ISOLATED POLYNUCLEOTIDES AND POLYPEPTIDES RELATING TO LOCI UNDERLYING RESISTANCE TO SOYBEAN CYST NEMATODE AND SOYBEAN SUDDEN DEATH SYNDROME AND METHODS EMPLOYING SAME

APPLICATION FOR
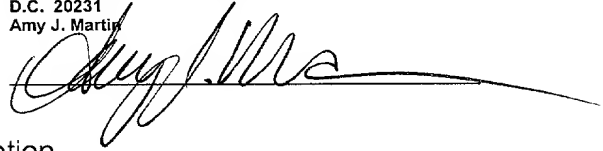
UNITED STATES LETTERS PATENT

By

David A. Lightfoot

Carbondale, Illinois

Khalid Meksem

Carbondale, Illinois

Description

ISOLATED POLYNUCLEOTIDES AND POLYPEPTIDES RELATING TO LOCI UNDERLYING RESISTANCE TO SOYBEAN CYST NEMATODE AND SOYBEAN SUDDEN DEATH SYNDROME AND METHODS EMPLOYING SAME

Cross Reference to Related Applications

This application is based on and claims priority to United States Provisional Application Serial Number 60/178,811, filed January 28, 2000, herein incorporated by reference in its entirety.

Technical Field

The present invention relates to plant breeding and plant genetics. More particularly, the invention relates to soybean cyst nematode and soybean sudden death syndrome resistance genes, soybean cyst nematode and soybean sudden death syndrome resistant soybean lines, and methods of breeding and engineering the same.

Table of Abbreviations

AFLP   -   amplified fragment length polymorphism

BAC   -   bacterial artificial chromosome

	bp	-	base pair
	Cf	-	tomato genes for resistance to *Cladosporium fulvus*
	FAM	-	6-carboxyfluorescein
5	FI	-	female index of parasitism
	indel	-	a nucleotide insertion or deletion
	MMAS	-	molecular marker-assisted selection
	QTL	-	quantitative trait loci
	RAPD	-	random amplified polymorphic DNA
10	RFLP	-	restriction fragment length polymorphism
	rhg1 and *Rhg4*	-	genetic loci conferring resistance to *Heterodera glycines*
	RIL	-	recombinant inbred line
	SCN	-	soybean cyst nematode
15	SDS	-	sudden death syndrome
	SSR	-	microsatellite
	TAMRA	-	6-carboxy- N,N,N'5N' tetrachlorofluorescein
	TET	-	6-carboxy-4,7,2',7',
20			tetrachlorofluorescein

Background of the Invention

Soybeans are a major cash crop and investment commodity in North America and elsewhere. Soybean oil is one of the most widely used edible oils, and soybeans are used worldwide both in animal feed and in human food production.

The soybean cyst nematode (SCN), *Heterodera glycines*, is a widespread pest of soybeans in the American continent. Reported first in Japan more than 75 years ago, since the first reports in North Carolina in 1954, SCN continues its spread toward almost all soybean-cultivated soils. Known as a small plant-parasitic roundworm that attacks the roots of soybeans, it reproduces very quickly, survives in the soil for many years in the absence of a soybean crop, and can cause substantial soybean crop yield losses.

Resistant soybean varieties are an effective tool available for SCN management. There are multiple sources for soybean cyst nematode resistance genes in commercial soybean varieties (PI88788, Peking and PI209332), and several have been used to develop cultivars (Myers & Anand (1991), *Euphytica* 55:197-201; Rao-Arrelli et al. (1988) *Crop Sci* 28:650-652). All the described loci involved in the resistance to SCN are reported to be quantitative. (Concibido et al. (1997) *Crop Sci* 37:258-264; Concibido (1996) *Theor Appl Genet* 93:234-241; Webb et al. (1995) *Theor Appl Genet* 91:574-581; Rao-Arrelli et al. (1992) *Crop Sci* 32:862-864; Matthews et al. (1991) *Soybean Genetics Newsletter*; Rao-Arrelli et al.,1988). They differ by their chromosomal position (LG A2, G, B, I, F, J and E) and race of the pathogen against which they confer the resistance (e.g. Race 1, 3, 5 or 14). SCN

resistance is simply inherited, but field resistance is oligogenic due to the existence of variation among SCN populations that are described as "races" (Riggs and Schmidt (1988) *J Nematol* 20:392-395).

One gene, *rhg1*, provides the major portion of resistance to SCN race 3
across many genotypes derived from Peking (Chang et al. (1997) *Crop Sci* 372:965-971; Mathews et al. (1998) *Theor Appl Genet* 97:1047-1052; Mahalingam et al. (1995) *Breed Sci* 45:435-445); PI437654 (Prabhu et al. (1999) *Crop Sci* 39:982-987; Webb et al.,1995), >PI88788= (Bell-Johnson et al. (1998) *Soybean Genet Newslett* 25:115-118; Concibido et al., 1997; Cregan
et al. (1999a) *Crop Sci* 39:1464-1490; Cregan et al. (1999b) *Theor Appl Genet* 99:811-818; Cregan et al. (1999c) *Theor Appl Genet* 99:918-928), >PI209332= (Concibido et al.,1996) , or >PI90763= (Concibido et al.,1997). A second gene for SCN resistance, *Rhg4*, provides an equal portion of resistance to SCN race 3 across genotypes derived from Peking (Chang et al., 1997; Mathews et al.,
1998; Mahalingam et al., 1995); and PI437654 (Prabhu et al., 1999; Webb et al., 1995) but not PI88788, PI209332 or PI90763 (Concibido et al., 1996; Concibido et al., 1997). Cytological studies suggest PI437654 and Peking derived resistances share mechanisms (pronounced necrosis and cell wall appositions) not seen in PI88788 in response to race 3 (Mahalingham et al.
(1996) *Genome* 39:986-998). These differences in mechanism may derive from distinct alleles at *Rhg4, rhg1* and/or other defense associated loci.

DNA molecular markers linked to SCN/SDS resistance loci can be used to develop effective plant breeding strategies. In general, molecular markers are abundant, often co-dominant, and suitable for rapid screening at the

seedling stage. Genetic linkage maps of soybean based on RFLP, RAPD, AFLP, and microsatellite markers have been described. See Brown et al. (1987) Principles and Practice of Nematode Control in Crops, pp179-232, Academic Press, Orlando Florida; Concibido et al., 1996; Concibido et al.,

5   1997; Mahalingham et al., 1995; Meksem et al. (1999) *Theor Appl Genet* 99:1131-1142; Meksem et al. (2000) *Theor Appl Genet* 101: 747-755; Webb et al., 1995; Weiseman et al. (1992) *Theor Appl Genet* 85:136-138; Lark et al. (1993) *Theor Appl Genet* 86:901-906; Shoemaker and Specht (1995) *Crop Sci* 35:436-446; Chang et al., 1997; Keim et al. (1997) *Crop Sci* 37:537-543).

10   All such markers have a limit of resistance trait predictability based principally on proximity of the marker to the resistance locus. In some cases, the interpretative value of genetic linkage experiments can be augmented through the simultaneous or serial detection of more than one genetic marker, although this also incurs additional time and resources. Thus, there is a need

15   for a reliable cost-effective method for detecting SCN or SDS resistance using genetic markers. Optimally, a genetic marker comprises a resistance gene.

Therefore, it is of particular importance, both to the soybean breeders and to farmers, to identify, genetic loci for resistance to SCN and SDS. Having knowledge of the loci for resistance to SCN and SDS, those of ordinary skill in

20   the art can breed or engineer SCN and SDS resistant soybeans. Soybean resistance can be further provided to a non-resistant cultivar in combination with other genotypic and phenotypic characteristics required for commercial soybean lines.

25

Summary of the Invention

The present invention discloses an isolated and purified genetic marker associated with SCN/SDS resistance in soybeans, said marker mapping to linkage group G in the soybean genome. Preferably, the marker has a sequence identical to any one of SEQ ID NOs:1, 3, and 5. Representative corresponding markers associated with SCN/SDS susceptibility are set forth as SEQ ID NOs:2, 4, and 6.

Also disclosed is an isolated and purified genetic marker associated with SCN/SDS resistance in soybeans, said marker mapping to linkage group A2 in the soybean genome. Preferably, the marker has a sequence identical to any one of SEQ ID NOs:7, 9, and 11. Representative corresponding markers associated with SCN/SDS susceptibility are set forth as SEQ ID NOs:8, 10, and 12.

The present invention further provides a plant, or parts thereof, which evidences an SCN/SDS resistance response comprising a genome, homozygous with respect to genetic alleles which are native to a first parent and nonnative to a second parent of the plant, wherein said second parent evidences significantly less resistant response to SCN/SDS than said first parent and said improved plant comprises alleles from said first parent that evidences resistance to SCN/SDS in hybrid combination in at least one locus selected from: a locus mapping to linkage group G and mapped by one or more of the markers set forth as SEQ ID NOs:1, 3, and 5, a locus mapping to linkage group A2 and mapped by one or more of the markers set forth as SEQ ID NOs:7, 9, and 11; or combinations thereof, said resistance not significantly

less than that of the first parent in the same hybrid combination, and yield characteristics which are not significantly different than those of the second parent in the same hybrid combination.

In another embodiment, a plant of the present invention, or parts thereof,

5   comprises the progeny of a cross between first and second inbred lines, alleles conferring SCN/SDS resistance being present in the homozygous state in the genome of one or the other or both of said first and second inbred lines such that the genome of said first and second inbreds together donate to the hybrid a complement of alleles necessary to confer the SCN/SDS resistance. Further

10   disclosed are hybrid plants derived therefrom.

Also disclosed herein are isolated and purified biologically active SCN/SDS resistance polypeptide and an isolated and purified nucleic acid molecule encoding the same are disclosed. Preferably, the polypeptide comprises a soybean SCN/SDS resistance polypeptide. Chimeric genes

15   comprising the isolated and purified nucleic acid molecules encoding a SCN/SDS resistance polypeptide are also provided.

In one embodiment, the nucleic acid molecule encoding a SCN/SDS resistance gene comprises an isolated soybean *rhg1* gene that confers SCN/SDS resistance to a non-resistant host organism. The gene is capable of

20   conveying *Heterodera glycines*-infestation resistance, *Fusarium solani*-infection resistance, or both *Heterodera glycines*-infestation resistance or *Fusarium solani*-infection resistance to a non-resistant plant germplasm, the gene located within a quantitative trait locus mapping to linkage group G and mapped by genetic markers of SEQ ID NOs:1, 3, and 5, said gene located along said

quantitative trait locus between said markers. Preferably, the polypeptide comprises (a) a polypeptide encoded by a nucleic acid sequence set forth as SEQ ID NO:13; (b) a polypeptide encoded by a nucleic acid having homology to a DNA sequence set forth as SEQ ID NO:13; (c) a polypeptide encoded by a

5   nucleic acid capable of hybridizing under stringent conditions to a nucleic acid comprising a sequence or the complement of a sequence set forth as SEQ ID NO:13; (d) a polypeptide which is a biologically functional equivalent of a peptide set forth as SEQ ID NO:14; or (e) a polypeptide comprising a fragment of a polypeptide of (a), (b), (c) or (d).

10   In another embodiment, the nucleic acid molecule encoding a SCN resistance polypeptide comprises an isolated soybean *Rhg4* gene that is capable of conveying *Heterodera glycines*-infestation resistance to a non-resistant plant germplasm, said gene located within a quantitative trait locus mapping to linkage group A2 and mapped by the AFLP markers of SEQ ID

15   NOs:7, 9, and 11, said gene located along said quantitative trait locus between said markers. Preferably, the nucleic acid molecule comprises any one of SEQ ID NOs:16-19.

The present invention further provides an isolated SCN/SDS resistance gene promoter region, or functional portion thereof, comprising an about 90 kb

20   fragment of soybean genomic clone 73P6 between *Bam*HI restriction sites and 21d9 between *Hin*DIII restriction site. The genomic clone is available from the Forrest BAC library described in Meksem et al (2000) *Theor Appl Genet* 101 5/6:747-755, available through Southern Illinois University-Carbondale (Carbondale, Illinois), Texas A&M University BAC center (College Station,

Texas), and Research Genetics (Huntsville, Alabama). Preferably, the isolated promoter region comprises the nucleotide sequence of SEQ ID NO:15 or a sequence substantially similar to SEQ ID NO:15. The SCN/SDS resistance gene promoter region can be operably linked to heterologous sequence.

5   A recombinant host cell comprising an isolated and purified nucleic acid molecule of the present invention is also disclosed, as is a transgenic plant having incorporated into its genome an isolated and purified nucleic acid molecule. In one embodiment, the nucleic acid molecule comprises encodes a SCN/SDS resistance polypeptide and is present in said genome in a copy

10   number effective to confer expression in the plant of the SCN/SDS resistance polypeptide. Seeds, parts or progeny of the transgenic plant are also disclosed.

Further provided is a method for detecting a nucleic acid molecule that encodes an SCN/SDS resistance polypeptide in a biological sample comprising nucleic acid material is also disclosed. The method comprises: (a) hybridizing

15   an isolated and purified nucleic acid molecule of the present invention under stringent hybridization conditions to the nucleic acid material of the biological sample, thereby forming a hybridization duplex; and (b) detecting the hybridization duplex. Preferably, the isolated and purified nucleic acid molecule comprises any of SEQ ID NOs:13 and 16-19.

20   An assay kit for detecting the presence, in biological samples, of an SCN/SDS resistance polypeptide is also disclosed. In one embodiment, the kit comprises a first container that contains a nucleic acid probe identical or complementary to a segment of at least ten contiguous nucleotide bases of a nucleic acid molecule of the present invention, preferably a nucleotide

sequence of any one of SEQ ID NOs:13 and 16-19. In another embodiment, the kit comprises a nucleic acid probe or primer identical to any one of SEQ ID NOs:1, 3, 5, 7, 9, and 11, or portion thereof.

A method for identifying soybean sudden death syndrome (SDS) resistance or soybean cyst nematode (SCN) resistance in a soybean plant using a SDS resistance gene, a SCN resistance gene, or DNA segments having homology to a SDS resistance gene or to an SCN resistance gene is also disclosed. In one embodiment, the method comprises: (a) probing nucleic acids obtained from the soybean plant with a probe derived from said SDS resistance gene or from said SCN resistance gene or from said DNA segment having homology to said SDS resistance gene or to said SCN resistance gene; and observing hybridization of said probe to said nucleic acids, the presence of said hybridization indicating SDS or SCN resistance in said soybean plant. In another embodiment, the method comprises (a) detecting a molecular marker linked to a quantitative trait locus associated with SCN/SDS resistance, wherein the molecular marker is the sequence set forth as any one of SEQ ID NOs:1, 3, 5, 7, 9, and 11; and (b) determining the presence of SCN/SDS resistance as detection of the molecular marker and determining the absence of SCN/SDS resistance as failure to detect the molecular marker of (b).

A method of reliably and predictably introgressing SCN/SDS resistance genes into non-resistant soybean germplasm is also disclosed. The method comprises: using one or more nucleic acid markers for marker assisted selection among soybean lines to be used in a soybean breeding program, wherein the nucleic acid markers map to linkage groups G or A2 and wherein

the nucleic acid markers are selected from among any of SEQ ID NOs: 1, 3, 5, 7, 9, and 11; and introgressing said resistance gene into said non-resistant soybean germplasm.

A soybean plant, or parts thereof, which evidences a SCN/SDS resistance response is also disclosed. The plant comprises a genome, homozygous with respect to genetic alleles which are native to a first parent and non-native to a second parent of the soybean plant, wherein said second parent evidences significantly less resistant response to SCN/SDS than said first parent, and said improved plant comprises alleles from said first parent that evidences resistance to SCN/SDS in hybrid combination of at least one locus selected from: a locus mapping to linkage group G and mapped by one or more of the markers set forth as SEQ ID NOs:1, 3, and 5, a locus mapping to linkage group A2 and mapped by one or more of the markers set forth in SEQ ID NOs:7, 9, and 11; or combinations thereof, said resistance not significantly less than that of the first parent in the same hybrid combination, and yield characteristics which are not significantly different than those of the second parent in the same hybrid combination.

The soybean plant, or parts thereof, can further comprise the progeny of a cross between first and second inbred lines, alleles conferring SCN/SDS resistance being present in a homozygous state in the genome of one or the other or both of said first and second inbred lines such that the genome of said first and second inbreds together donate to the hybrid a complement of alleles necessary to confer the SCN/SDS resistance. Thus, an SCN/SDS resistant hybrid, or parts thereof, formed with the soybean plant is also disclosed, as is a

soybean plant, or parts thereof, formed by selfing the SCN/SDS resistant hybrid.

A method of positional cloning of a nucleic acid is also disclosed. The the method comprises: (a) identifying a first nucleic acid genetically linked to a

5 SCN/SDS resistance locus, wherein the first nucleic acid maps between two markers selected from SEQ ID NOs:1-12; and (b) cloning the first nucleic acid. Optionally, the first nucleic acid can comprise the *rhg1* locus or the *Rhg4* locus.

A method for producing an antibody that specifically recognizes a SCN/SDS resistance polypeptide is also disclosed. The method comprises (a)

10 recombinantly or synthetically producing a SCN/SDS resistance polypeptide, or portion thereof; (b) formulating the polypeptide of (a) whereby it is an effective immunogen; (c) administering to an animal the formulation of (b) to generate an immune response in the animal comprising production of antibodies, wherein antibodies are present in the blood serum of the animal; and (d) collecting the

15 blood serum from the animal of (c) comprising antibodies that specifically recognize a SCN/SDS resistance polypeptide. Also provided is an antibody produced by the disclosed method.

Methods for identifying a candidate compound as a modulator of SCN/SDS resistance activity is also disclosed. Such methods include but are

20 not limited to cell-based assays of SCN/SDS resistance gene expression, assays of specific binding to SCN/SDS regulatory elements, and assays of specific binding to SCN/SDS polypeptides. Optionally, the screening methods are adapted to a high-throughput format.

In one embodiment, the method comprises: (a) exposing a cell sample with a candidate compound to be tested, the cell sample containing at least one cell containing a DNA construct comprising a modulatable transcriptional regulatory sequence of an SCN/SDS resistance-encoding nucleic acid and a reporter gene which is capable of producing a detectable signal; (b) evaluating an amount of signal produced in relation to a control sample; and (c) identifying a candidate compound as a modulator of SCN/SDS resistance activity based on the amount of signal produced in relation to a control sample.

The present invention also provides a method for identifying a substance that regulates SCN/SDS resistance gene expression using a chimeric gene that includes an isolated SCN/SDS resistance gene promoter region operably linked to a reporter gene. According to this method, a gene expression system is established that includes the chimeric gene and components required for gene transcription and translation so that reporter gene expression is assayable. To select a substance that regulates SCN/SDS resistance gene expression, the method further provides the steps of using the gene expression system to determine a baseline level of reporter gene expression in the absence of a candidate regulator; providing a plurality of candidate regulators to the gene expression system; and assaying a level of reporter gene expression in the presence of a candidate regulator. A candidate regulator is selected whose presence results in an altered level of reporter gene expression when compared to the baseline level. Preferably, the isolated SCN/SDS resistance gene promoter region used in this method comprises the sequence of SEQ ID NO:15, or functional portion thereof.

In another embodiment, the method comprises using an SCN/SDS regulatory sequence to identify a candidate substance that specifically binds to the regulatory sequence. According to the method, a SCN/SDS regulatory gene sequence is exposed to a candidate substance under conditions suitable

5   for binding to a nucleic acid sequence, and a candidate regulator is selected that specifically binds to the SCN/SDS resistance gene promoter region. Preferably, the isolated SCN/SDS resistance gene promoter region used in this method comprises the sequence of SEQ ID NO:15, or functional portion thereof.

10   In another embodiment, a cell-free assay system is used and comprises: (a) exposing a SCN/SDS polypeptide of the present invention to a candidate compound; (b) assaying binding of the candidate compound to the SCN/SDS polypeptide; and (c) identifying a candidate compound as a putative modulator of SCN/SDS resistance activity based on specific binding of the candidate

15   compound to the SCN/SDS polypeptide. Preferably, the SCN/SDS polypeptide comprises some or all of the amino acids of SEQ ID NO:14.

A method of modulating SCN/SDS resistance in a plant is also disclosed. The method comprises administering to the plant an effective amount of a substance that modulates expression of an SCN/SDS resistance activity-

20   encoding nucleic acid molecule in the plant to thereby modulate SCN/SDS resistance in the plant. Preferably, the substance that modulates expression of an SCN/SDS resistance activity is discovered by a disclosed method of the present invention.

A method for providing a resistance characteristic to a plant is also disclosed. The method comprises introducing to said plant a construct comprising a nucleic acid sequence encoding an SCN/SDS resistance gene product operatively linked to a promoter, wherein production of the SCN/SDS

5   resistance gene product in the plant provides a resistance characteristic to the plant. The construct can further comprises a vector selected from the group consisting of a plasmid vector or a viral vector. The SCN/SDS resistance gene product comprises a protein having an amino acid sequence of SEQ ID NO:14. The nucleic acid sequence comprises the nucleotide sequence of SEQ ID

10   NO:13 or a nucleic acid that is substantially similar to SEQ ID NO:13, and which encodes an SCN/SDS resistance polypeptide.

The resistance characteristic is preferably nematode resistance, fungal resistance or combinations thereof. More preferably, the nematode resistance is *H. glycines* resistance, even more preferably race 3 *H. glycines* resistance.

15   In an alternative embodiment the construct further comprises another nucleic acid molecule encoding a polypeptide that provides an additional desired characteristic to the plant. Optionally, the method further comprises monitoring an insertion point for the construct in the plant genome; and providing for insertion of the construct into the plant genome at a location not

20   associated with the resistance characteristic, the desired characteristic, or both the resistance and the desired characteristic. Preferably, the plant is a soybean plant.

The present invention also provides methods for providing a resistance characteristic to a plant is also disclosed, wherein a combination of genetic and

non-genetic techniques is employed. The method comprises introducing to said plant a construct comprising a nucleic acid sequence encoding an SCN/SDS resistance gene product operatively linked to a promoter and provision of a substance that modulates SCS/SDS resistance gene activity,

5   wherein production of the SCN/SDS resistance gene product in the plant, in combination with provision of the SCN/SDS resistance gene modulator, provides a resistance characteristic to the plant.

Accordingly, it is an object of the present invention to provide novel isolated polynucleotides and polypeptides relating to loci underlying resistance

10   to soybean cyst nematode and soybean sudden death syndrome and methods employing same. The object is achieved in whole or in part by the present invention.

An object of the invention having been stated hereinabove, other objects and advantages will become evident as the description proceeds, when taken

15   in connection with the accompanying Drawings and Examples as best described hereinbelow.

Brief Description of the Drawings

Figure 1 depicts new AFLP genetic markers for SCN/SDS resistance.

20   Figure 1A presents genomic sequences of the both alleles (resistant Forrest and susceptible Essex) of the converted AFLP markers $E_{ATG}M_{CGA}87$ (SEQ ID NOs:1-2); $E_{CTA}M_{AGG}113$ (SEQ ID NOs:3-4); $E_{CGG}M_{AGA}116$ (SEQ ID NOs:5-6); $E_{CCG}M_{AAC}405$ (SEQ ID NOs:7-8), $E_{CCC}M_{ATG}161$ (SEQ ID NOs:9-10), $E_{CCA}M_{AGC}114$ (SEQ ID NOs:11-12. The italicized and underlined sequences

represent the forward and reverse sequence specific primers used. The bold capital sequences represent the original AFLP restriction site. The bold letters indicate the difference in sequences between the two alleles.

Figure 1B presents genomic sequences of the two alleles (resistant and susceptible) of the converted $E_{ATG}M_{CGA}87$ markers. The italic sequences represent the resistance specific TaqMan™ probes TMA5-RE and the susceptible allele specific probe TMA5-S. The standard font underlined sequence represent the TaqMan™ forward and reverse primers assay, the underlined italic sequence is the ATG4BACF primer used for sequence extension of the $E_{ATG}M_{CGA}87$ marker, the BAC derived extended sequences are in small font capitals.

Figure 2 depicts AFLPs for selecting SCN/SDS resistance.

Figure 2A shows PCR amplification products using $E_{ATG}M_{CGA}87$ sequence specific primers TMA5 forward and reverse: Lane 1-40 represent 40 RIL DNA, 41 and 42 are the two parents. F: Forrest; E: Essex; 1: resistant allele; 2: susceptible allele; H: heterozygote lines. The PCR products were separated by electrophoresis on a 4% (w/v) Metaphor gel.

Figure 2B shows a partial AFLP autoradiograph profile of the $E_{CGG}M_{AGA}116$ marker. The six selective nucleotides step was replaced by *Mse*I primer $M_{AGAGACT}$ and *Eco*RI primer E. Lane 7: Essex; Lane 8: Forrest; Lane 1 to 6 and 9 to 20 represent RIL DNA; 1: resistant allele; 2: susceptible allele

Figure 2C shows PCR amplification products using $E_{CTA}M_{AGG}113$ sequence specific primers CTA forward and reverse: Lane 1-40 represent 40 RIL DNA, 41 and 42 are the two parent. F: Forrest; E: Essex; 1: resistant allele;

2: susceptible allele; H: heterozygote lines. The PCR products were separated by electrophoresis on a 4% (w/v) Metaphor gel.

Figure 2D shows PCR amplification products using $E_{CCG}M_{AAC}405$ sequence specific primers A2D8 forward and reverse: Lane 1-40 represent 40

5   RIL DNA, 41 and 42 are the two parents F: Forrest; E: Essex; 1: resistant allele; 2: susceptible allele; H: heterozygote lines. The PCR products were separated by electrophoresis on a 4% (w/v) Metaphor gel.

Figure 3 depicts a genetic and physical map showing the location of an *Rhg4* gene relative to DNA markers. The location of the aspartokinase serine

10   dehydrogenase (AK-HSDH) and the A2D8 marker are indicated as determined by restriction mapping of BAC DNA. The A2D8 sequences for Essex and Forrest alleles are deposited in GenBank as Accession Nos. AF286701 and AF286700, respectively. The *I* locus (I) position was estimated by relation to BARC-SAT_162 (Cregan et al., 1999c). Genetic mapping shows *Rhg4* and

15   A2D8 are both within the interval shown by the horizontal line and within a large insert clone, 100B10, that contains a 140 kbp insert (Zobrist et al. (2000) *Soybean Genet Newslett* 27:10-15).

Figure 4 depicts the gene structure of the *rhg1* gene and clones derived from Forrest genomic DNA.

20   Figure 5 depicts detection of the A2D8 marker polymorphism using the TaqMan™ assay and manual selection of genotypes. Eighty-six individuals from an F5 derived population of recombinant inbred lines from the cross of Essex x Forrest that segregate for resistance to SCN are shown.

Figure 5A is an image of fluorescent signals viewed under the "dye component" field of the sequence detection software and the A2D8 genotypes were manually selected based on the ratio of FAM and TET signals. Allele 1 homozygous, Forrest type; FAM<< TET. Allele 2 homozygous, Essex type;

5   TET<< FAM. Alleles 1 and 2 heterogeneous, Essex and Forrest type; TET less than 2 fold greater or lesser than FAM. Two selections were used, in the first (TaqManTM assay1) group of genotypes FAM 6-8 and TET 8-9 were considered susceptible. In the second (TaqMan™ assay 2) group, they were considered heterogeneous.

10   Figure 5B is a spreadsheet that contains scores (allele designations) for the samples as they were arranged in the 96 well plate. There was no DNA in wells E12, F12 and G12 (negative controls). There was Essex DNA in wells A1, C12 and D12. There was Forrest DNA in wells B2, A12 and B12. The RIL DNA was in well A3 to H11 in order by row from RIL1-RIL86 except samples E1

15   (RIL3) and E6 (RIL 43) that did not amplify. The RILs resistant to SCN had an index of parasitism FI <10% of the susceptible check resistant lines.

Figure 6 depicts detection of the A2D8 marker polymorphism by PCR amplification and gel electrophoresis of soybean genotypes. Seventy-eight individuals from an F5 derived population of recombinant inbred lines from the

20   cross of Essex x Forrest that segregate for resistance to SCN are shown.

Figure 6A is an image of fluorescent signals viewed under the "dye component" field of the sequence detection software and the A2D8 genotypes were manually selected based on the ratio of FAM and TET signals. Lane 1, 42 Essex; Lane 2 and 41 Forrest; Lanes 3-40 RILS 1-38.

Figure 6B is a picture of an ehtidium-stained gel, showing resolution of gel electrophoresis markers. Lane 42 Essex; Lane 41 Forrest; Lanes 1-40 RILS 39-78. Asterisks indicate disagreements with the TaqMan™ assay 1.

Figure 7A-B presents the *rhg1* gene seqeunce (SEQ ID NO:13).

5   Figure 7C presents the rhg1 polypeptide (SEQ ID NO:14).

Figure 7D shows sequences producing significant alignments using BLAST analysis.

Figure 7E-F is an alignment between rhg1 protein (SEQ ID NO:14) and *Arabidopsis thaliana* hypothetical protein T18N14.120 (GenBank Accession

10   T46070).

Detailed Description of the Invention

Disclosed herein is the identification of AFLP markers that are genetically linked to the SCN/SDS resistance loci of Forrest. Further disclosed

15   are purified and isolated SCN or SDS resistance genes, proximal sequences to SCN/SDS resistance genes, and SCN/SDS resistance-related genes.

The isolated and purified polynucleotide sequences disclosed herein can thus be used in a variety of applications pertaining to breeding and engineering soybeans having SCN and SDS resistance. For example, the isolated

20   polynucleotides disclosed herein can be used in position-based or homology-based cloning of additional SCN/SDS resistance genes, including regulatory elements; in gene structure determination; in studies of genome organization and gene expression; in gene complementation experiments; in the isolation of

additional DNA markers for gene manipulation and molecular marker assisted breeding; and in plant transformation and the production of transgenic plants.

The present invention also pertains to a soybean plant and methods of producing the same, which is resistant to soybean cyst nematodes (SCN). In one embodiment, the method comprises stable transformation of a plant with an *rhg1* gene, disclosed herein. In another embodiment, the method comprises introgression in soybean of a trait enabling the plant to resist soybean cyst nematode (SCN) infestation. Additionally, the present invention relates to method of precise and accurate introgression of the genetic material conferring SCN resistance from one or more parent plants into the progeny.

The present invention also pertains to a soybean plant and methods of producing the same, which is resistant to soybean sudden death syndrome (SDS). In one embodiment, the method comprises stable transformation of a plant with an *rhg1* gene, disclosed herein. In another embodiment, the method comprises introgression of the genetic material conferring SDS resistance from one or more parent plants into the progeny with precision and accuracy.

The invention differs from present technology in several regards. In one aspect, the present invention provides the first disclosure of the *rhg1* gene sequence, thereby enabling transgenic approaches for providing SCN/SDS resistance. Further, the present invention provides a non-electorphoretic selection assay using nucleotide sequences of SCN/SDS resistance gene alleles. The disclosed nucleotide sequences of SCN/SDS resistance genes and associated genetic markers provide means for easily selecting resistant cultivars, for assembling many resistance genes in a single cultivar, for

combining resistance genes in novel combinations, for identifying genes that confer resistance in new cultivars, and for predicting resistance in cultivars. The invention is used to improve selection for SDS and SCN resistance in soybean in breeding programs.

5   I.   <u>Traits</u>

The term "phenotype" or "trait" each refer to any observable property of an organism, produced by the interaction of the genotype of the organism and the environment. A phenotype can encompass variable expressivity and penetrance of the phenotype. Exemplary phenotypes include but are not

10   limited to a visible phenotype, a physiological phenotype, a susceptibility phenotype, a cellular phenotype, a molecular phenotype, and combinations thereof. Preferably, the phenotype is related to SCN/SDS resistance. The term "susceptibility phenotype" refers to an increased capacity or risk for displaying a phenotype, i.e. a susceptibility to SCN/SDS infection.

15   The term "complex trait" as used herein refers to a trait that is not inherited as predicted by classical Mendelian genetics. A complex trait results from the interaction of multiple genes, each gene contributing to the phenotype. Complex traits can be continuous or show threshold penetrance. In the field, SCN/SDS resistance is inherited as a complex trait.

20   The term "quantitative trait" is a complex trait that can be assessed quantitatively. Quantitation entails measurement of a trait across a continuous distribution of values. SCN/SDS resistance is a quantitative trait.

The term "SCN/SDS resistance" or "SCN/SDS resistance trait" as used herein refers to a cellular or organismal capacity for resistance to nematode or

fungal infection, or both. Preferably, the nematode resistance is *Heterodera glycines* (the organism that causes SCN in soybeans) resistance, even more preferably race 3 *Heterodera glycines* resistance. The fungal resistance is preferably *Fusarium solani* (the organism that causes SDS in soybeans)-

5   infection resistance. SCN resistance can be assayed in the field or in the greenhouse by methods known in the art, including but not limited to determination of an SCN index of parasitism as disclosed in Example 2, Meksem et al. (1999), and U.S. Patent No. 6,096,944. SDS resistance can be scored by determination of disease incidence, disease severity, and disease

10   index values as disclosed in Hnetkovsky et al. (1996) Crop Sci 36(2):393-400, Njiti et al. (1996) *Crop Sci* 36:1165-1170; and Matthews et al. (1991).

The term "SCN/SDS resistance" is used herein for convenience to describe traits, transgenic plants, polynucleotides, and polypeptides of the present invention. Therefore, the resistance characteristic conveyed by the

15   polynucleotides and polypeptides of the present invention refers to any resistance characteristic as set forth herein and as would be apparent to one of ordinary skill in the art after reviewing the disclosure of the present invention.

The term "molecular phenotype" refers to a detectable feature of molecules in a cell or organism. Exemplary molecular phenotypes include but

20   are not limited to a presence of a genetic marker nucleotide sequence, a presence of a SCN/SDS resistance gene sequence, a level of gene expression, a splice selection, a level of protein, a protein type, a protein modification, a level of lipid, a lipid type, a lipid modification, a level of carbohydrate, a carbohydrate type, a carbohydrate modification, and combinations thereof.

Methods for observing, detecting, and quantitating molecular phenotypes are well known to one skilled in the art. See Sambrook et al., eds. (1989) Molecular Cloning, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, New York; by Silhavy et al. (1984) Experiments with Gene Fusions, Cold Spring

5   Harbor Laboratory Press, Cold Spring Harbor, New York, New York; by Ausubel et al. (1992) Current Protocols in Molecular Biology, John Wylie and Sons, Inc. New York, New York; Landgren et. al. (1988) Science 242:229-237; Bodanszky, et al. (1976) Peptide Synthesis, John Wiley and Sons, Second Edition, New York, New York; Harlow and Lane (1988) Antibodies: A

10   Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; Ochman et al. (1990) in PCR protocols: a Guide to Methods and Applications, Innis et al. (eds.), pp. 219-227, Academic Press, San Diego, California; Koduri and Poola (2001) Steroids 66(1):17-23; Regan et al. (2000) Anal Biochem 286(2):265-276; U.S. Patent Nos. 6,096,555; 5,958,624; and

15   5,629,158.

II.   Genetic Mapping

For genetic mapping, a representative population was generated as in Example 1. To detect genomic regions associated with resistance to SCN and resistance to SDS, the RILs were classified as Essex type or Forrest type for

20   each marker. In some cases, SCN susceptibility and resistance was quantitatively determined according to a SCN female index (F1) of parasitism (Meksem, 1999) as described in Example 2. Markers were compared with SCN or SDS response scores by the F-test in analysis of variance (ANOVA) done with SAS (SAS Institute Inc., Cary, North Carolina, 1988). The probability of

association of each marker with each trait was determined and a significant association was declared if $P \leq 0.05$ (unless noted otherwise in the text) since the detection of false associations is reduced in isogenic lines (Landers & Botstein (1989) *Genetics* 121:185-199; Paterson et al. (1990) *Genetics*

5   124:735-742).

Selected pairs of markers were analyzed by the two-way ANOVA using the general linear model (PROC GLM) procedure to detect non-additive interactions between the unlinked QTL (Chang et al. (1996) *Crop Sci* 36:965-971) or Epistat (Chase et al. (1997) *Theor Appl Genet* 94:724-730). Non-

10   additive interactions between markers which were significantly associated with SCN/SDS response were excluded when $P \geq 0.05$. Selected groups of markers were analyzed by multi-way ANOVA to estimate joint heritabilities for traits associated with multiple QTL. Joint heritability was determined from the R^2 term for the joint model in multi-way ANOVA.

15   Mapmaker-EXP 3.0 (Lander et al. 1987) was used to calculate map distances (cM, Haldane units) between linked markers and to construct a linkage map including traits as genes. The RIL (recombinant inbred line) and F_3 self genetic models were used. The $\log_{10}$ of the odds ratio (LOD) for grouping markers was set minimally at 2.0, and maximum distance was set at 30 cM.

20   Conflicts were resolved in favor of the highest LOD score after checking the raw data for errors. Marker order within groups was determined by comparing the likelihood of many map orders. A maximum likelihood map was computed with error detection. Trait data were used for QTL analysis (Webb et al. 1995; Chang et al. 1997). The data were subjected to ANOVA (SAS Institute Inc.,

Cary, North Carolina) with mean separation by LSD (Gomez and Gomez (1984). Graphs were constructed by Quattro Pro version 5.0 (Novell Inc., Orem, Utah).

III.   Nucleotide Sequences of SCN/SDS Resistance Genes and Associated

5   Genetic Markers

The nucleic acid molecules provided by the present invention include the isolated nucleic acid molecules of SEQ ID NOs:1-13 and 15-114, sequences substantially similar to sequences of SEQ ID NOs:1-13 and 15-114, conservative variants thereof, plant-expressible variants thereof, subsequences

10   and elongated sequences thereof, complementary DNA molecules, and corresponding RNA molecules. The present invention also encompasses genes, cDNAs, promoters, chimeric genes, and vectors comprising disclosed SCN/SDS resistance gene and SCN/SDS resistance gene marker nucleic acid sequences.

15   III.A.   General Considerations

The term "nucleic acid molecule" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides that have similar properties as the

20   reference natural nucleic acid. Unless otherwise indicated, a particular nucleotide sequence also implicitly encompasses conservatively modified variants thereof *(e.g.* degenerate codon substitutions), complementary sequences, subsequences, elongated sequences, as well as the sequence explicitly indicated. The terms "nucleic acid molecule" or "nucleotide sequence"

can also be used in place of "gene", "cDNA", or "mRNA". Nucleic acids can be derived from any source, including any organism.

The term "isolated", as used in the context of a nucleic acid molecule, indicates that the nucleic acid molecule exists apart from its native environment and is not a product of nature. An isolated DNA molecule can exist in a purified form or can exist in a non-native environment such as a transgenic host cell.

The term "purified", when applied to a nucleic acid, denotes that the nucleic acid is essentially free of other cellular components with which it is associated in the natural state. Preferably, a purified nucleic acid molecule is a homogeneous dry or aqueous solution. The term "purified" denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Particularly, it means that the nucleic acid is at least about 50% pure, more preferably at least about 85% pure, and most preferably at least about 99% pure.

The term "substantially identical", in the context of two nucleotide or amino acid sequences, can also be defined as two or more sequences or subsequences that have at least 60%, preferably 80%, more preferably 90-95%, and most preferably at least 99% nucleotide or amino acid sequence identity, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms (described herein below under the heading Nucleotide and Amino Acid Sequence Comparisons) or by visual inspection. Preferably, the substantial identity exists in nucleotide sequences of at least 50 residues, more preferably in nucleotide sequence of at least about 100 residues, more preferably in

nucleotide sequences of at least about 150 residues, and most preferably in nucleotide sequences comprising complete coding sequences.

In one aspect, polymorphic sequences can be substantially identical sequences. The term "polymorphic" refers to the occurrence of two or more

5   genetically determined alternative sequences or alleles in a population. An allelic difference can be as small as one base pair.

Another indication that two nucleotide sequences are substantially identical is that the two molecules specifically or substantially hybridize to each other under stringent conditions. In the context of nucleic acid hybridization,

10   two nucleic acid sequences being compared can be designated a "probe" and a "target". A "probe" is a reference nucleic acid molecule, and a "target" is a test nucleic acid molecule, often found within a heterogenous population of nucleic acid molecules. "Target sequence" is synonymous with "test sequence".

A preferred nucleotide sequence employed for hybridization studies or

15   assays includes probe sequences that are complementary to or mimic at least an about 14 to 40 nucleotide sequence of a nucleic acid molecule of the present invention. Preferably, a probe comprises 14 to 20 nucleotides, or even longer where desired, such as 30, 40, 50, 60, 100, 200, 300, or 500 nucleotides or up to the full length of any of SEQ ID NOs:1-13, 15-114. Such fragments

20   can be readily prepared by, for example, directly synthesizing the fragment by chemical synthesis, by application of nucleic acid amplification technology, or by introducing selected sequences into recombinant vectors for recombinant production. The phrase "hybridizing specifically to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence

under stringent conditions when that sequence is present in a complex nucleic acid mixture (e.g., total cellular DNA or RNA). The phrase "binds substantially to" refers to complementary hybridization between a probe nucleic acid molecule and a target nucleic acid molecule and embraces minor mismatches

5   that can be accommodated by reducing the stringency of the hybridization media to achieve the desired hybridization. Probe sequences can also hybridize specifically to duplex DNA under certain conditions to form triplex or other higher order DNA complexes. The preparation of such probes and suitable hybridization conditions are well known in the art.

10   "Stringent hybridization conditions" and "stringent hybridization wash conditions" in the context of nucleic acid hybridization experiments such as Southern and Northern blot analysis are both sequence- and environment-dependent. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen

15   (1993) <u>Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes</u> part I chapter 2, Elsevier, New York, New York. Generally, highly stringent hybridization and wash conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. Typically, under

20   "stringent conditions" a probe will hybridize specifically to its target subsequence, but to no other sequences.

The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Very stringent conditions are selected to be equal to the T_m for a particular probe.

An example of stringent hybridization conditions for Southern or Northern Blot analysis of complementary nucleic acids having more than about 100 complementary residues is overnight hybridization in 50% formamide with 1 mg of heparin at 42°C. An example of highly stringent wash conditions is 15

5 minutes in 0.15 M NaCl at 65°C. An example of stringent wash conditions is 15 minutes in 0.2X SSC buffer at 65°C (See Sambrook et al., 1989) for a description of SSC buffer). Often, a high stringency wash is preceded by a low stringency wash to remove background probe signal. An example of medium stringency wash conditions for a duplex of more than about 100 nucleotides, is

10 15 minutes in 1X SSC at 45°C. An example of low stringency wash for a duplex of more than about 100 nucleotides, is 15 minutes in 4-6X SSC at 40°C. For short probes (e.g., about 10 to 50 nucleotides), stringent conditions typically involve salt concentrations of less than about 1.0 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0-8.3, and the

15 temperature is typically at least about 30°C. Stringent conditions can also be achieved with the addition of destabilizing agents such as formamide. In general, a signal to noise ratio of 2-fold (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization.

20 The following are examples of hybridization and wash conditions that can be used to clone homologous nucleotide sequences that are substantially identical to reference nucleotide sequences of the present invention: a probe nucleotide sequence preferably hybridizes to a target nucleotide sequence in 7% sodium dodecyl sulfate (SDS), 0.5 M $NaPO_4$, 1 mM EDTA at 50°C followed

by washing in 2X SSC, 0.1% SDS at 50°C; more preferably, a probe and target sequence hybridize in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO$_4$, 1 mM EDTA at 50°C followed by washing in 1X SSC, 0.1% SDS at 50°C; more preferably, a probe and target sequence hybridize in 7% sodium dodecyl sulfate

5   (SDS), 0.5 M NaPO$_4$, 1 mM EDTA at 50°C followed by washing in 0.5X SSC, 0.1% SDS at 50°C; more preferably, a probe and target sequence hybridize in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO$_4$, 1 mM EDTA at 50°C followed by washing in 0.1X SSC, 0.1% SDS at 50°C; more preferably, a probe and target sequence hybridize in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO$_4$, 1

10   mM EDTA at 50°C followed by washing in 0.1X SSC, 0.1% SDS at 65°C.

A further indication that two nucleic acid sequences are substantially identical is that proteins encoded by the nucleic acids are substantially identical, share an overall three-dimensional structure, are biologically functional equivalents; or are immunologically cross-reactive. These terms are defined

15   further under the heading SCN/SDS Resistance Polypeptides herein below. Nucleic acid molecules that do not hybridize to each other under stringent conditions are still substantially identical if the corresponding proteins are substantially identical. This can occur, for example, when two nucleotide sequences are significantly degenerate as permitted by the genetic code.

20   The term "conservatively substituted variants" refers to nucleic acid sequences having degenerate codon substitutions wherein the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer et al. (1991) Nucleic Acid Res. 19:5081; Ohtsuka

et al. (1985) *J Biol Chem* 260:2605-2608; Rossolini et al. (1994) *Mol Cell Probes* 8:91-98).

The term "plant-expressible variant" means a substantially similar sequence that has been modified to comprise a coding sequence (nucleotide

5   sequence) can be efficiently expressed by plant cells, tissue and whole plants. The art understands that a plant-expressible coding sequence has a GC composition consistent with good gene expression in plant cells, a sufficiently low CpG content so that expression of that coding sequence is not restricted by plant cells, and codon usage which is consistent with that of plant genes.

10   Where it is desired that the properties of the plant-expressible SCN/SDS resistance gene are identical to those of the naturally occurring SCN/SDS resistance gene, the plant-expressible homolog will have an identical coding sequence or a substantially identical coding sequence.

The term "subsequence" refers to a sequence of nucleic acids that

15   comprises a part of a longer nucleic acid sequence.   An exemplary subsequence is a probe, described herein above, or a primer.   The term "primer" as used herein refers to a contiguous sequence comprising about 8 or more deoxyribonucleotides or ribonucleotides, preferably 10-20 nucleotides, and more preferably 20-30 nucleotides of a selected nucleic acid molecule.

20   The primers of the present invention encompass oligonucleotides of sufficient length and appropriate sequence so as to provide initiation of polymerization on a nucleic acid molecule of the present invention.

The term "elongated sequence" refers to an addition of nucleotides (or other analogous molecules) incorporated into the nucleic acid. For example, a polymerase (e.g., a DNA polymerase), e.g., a polymerase that adds sequences at the 3' terminus of the nucleic acid molecule can be employed to prepare an

5 elongated sequence. In addition, the nucleotide sequence can be combined with other DNA sequences, such as promoters, promoter regions, enhancers, polyadenylation signals, intronic sequences, additional restriction enzyme sites, multiple cloning sites, and other coding segments.

The term "complementary sequence", as used herein, indicates two

10 nucleotide sequences that comprise anti-parallel nucleotide sequences capable of pairing with one another upon formation of hydrogen bonds between base pairs. As used herein, the term "complementary sequences" means nucleotide sequences which are substantially complementary, as can be assessed by the same nucleotide comparison set forth above, or is defined as being capable of

15 hybridizing to the nucleic acid segment in question under relatively stringent conditions such as those described herein. A particular example of a complementary nucleic acid segment is an antisense oligonucleotide.

The present invention further includes vectors comprising the disclosed SCN/SDS resistance gene sequences, including plasmids, cosmids, and viral

20 vectors. The term "vector", as used herein refers to a DNA molecule having sequences that enable its replication in a compatible host cell. A vector also includes nucleotide sequences to permit ligation of nucleotide sequences within the vector, wherein such nucleotide sequences are also replicated in a

compatible host cell. A vector can also mediate recombinant production of an SCN/SDS resistance gene polypeptide, as described further herein below.

Nucleic acids of the present invention can be cloned, synthesized, recombinantly altered, mutagenized, or combinations thereof. Standard
5   recombinant DNA and molecular cloning techniques used to isolate nucleic acids are well known in the art. Exemplary, non-limiting methods are described by Sambrook et al., eds.,1989; by Silhavy et al., 1984; by Ausubel et al., 1992; and by Glover, ed. (1985) DNA Cloning: A Practical Approach, MRL Press, Ltd., Oxford, United Kingdom. Site-specific mutagenesis to create base pair
10   changes, deletions, or small insertions are also well known in the art as exemplified by publications, see, e.g., Adelman et al., (1983) DNA 2:183; Sambrook et al. (1989).

Nucleotide sequences of the present invention can detected, subcloned, sequenced, and further evaluated by any measure well known in the art using
15   any method usually applied to the detection of a specific DNA sequence including but not limited to dideoxy sequencing, PCR, oligomer restriction (Saiki et al., Bio/Technology 3:1008-1012 (1985), allele-specific oligonucleotide (ASO) probe analysis (Conner et al. (1983) Proc Natl Acad Sci USA 80:278), and oligonucleotide ligation assays (OLAs) (Landgren et. al. (1988) Science
20   241:1007). Molecular techniques for DNA analysis have been reviewed (Landgren et. al. (1988) Science 242:229-237).

Table of Functionally Equivalent Codons

Amino Acids			Codons
Alanine	Ala	A	GCA GCC GCG GCU
25 Cysteine	Cys	C	UGC UGU
Aspartic Acid	Asp	D	GAC GAU

	Glumatic acid	Glu	E	GAA GAG
	Phenylalanine	Phe	F	UUC UUU
	Glycine	Gly	G	GGA GGC GGG GGU
	Histidine	His	H	CAC CAU
5	Isoleucine	Ile	I	AUA AUC AUU
	Lysine	Lys	K	AAA AAG
	Leucine	Leu	L	UUA UUG CUA CUC CUG CUU
	Methionine	Met	M	AUG
	Asparagine	Asn	N	AAC AAU
10	Proline	Pro	P	CCA CCC CCG CCU
	Glutamine	Gln	Q	CAA CAG
	Arginine	Arg	R	AGA AGG CGA CGC CGG CGU
	Serine	Ser	S	ACG AGU UCA UCC UCG UCU
	Threonine	Thr	T	ACA ACC ACG ACU
15	Valine	Val	V	GUA GUC GUG GUU
	Tryptophan	Trp	W	UGG
	Tyrosine	Tyr	Y	UAC UAU

III.B. Genetic Markers

20   The term "genetic marker", as used herein generally refers to a genetic

locus, a phenotype conferred by locus, or a nucleotide sequence residing at a

locus, wherein the locus is genetically linked to a trait of interest. The term

"genetically linked" as used herein refers to two or more loci that are predictably

inherited together during random crossing or intercrossing. Quantitative linkage

25   analysis is further described in the section Genetic Mapping herein above.

Preferably, genetically linked loci are less than about 10 cM apart, more

preferably less than about 5 cM apart, and even more preferably less than

about 1cM apart. Optimally, the genetic marker and the gene conferring a trait

of interest comprise the same or overlapping nucleotide sequence.

30   An embodiment of the present invention comprises genetic markers

associated with SCN resistance and SDS resistance that are isolatable from

soybeans, and which are free from total genomic DNA. Disclosed herein are

sequences of AFLP markers mapped in soybean to the chromosomal segments

carrying *rhg1* and SDS loci on molecular linkage group G and the *Rhg4* locus on molecular linkage group A2. Representative markers for SCN/SDS resistance are set forth as SEQ ID NOs:1, 3, 5, 7, 9, and 11. Respresentative corresponding markers for SCN/SDS susceptibility are set forth as SEQ ID

5　　NOs:2, 4, 6, 8, 10, and 12.

　　　　AFLP bands were obtained as described in Example 3. From each AFLP band, 4-30 clones were sequenced (mean 15.6) depending on the sequence complexity of the originating band. The sequence analysis showed that each AFLP band can be composed of a number of different DNA

10　　sequences from fragments of identical size. A mean of 6 sequences per band with a range of 1-15 sequences per band was detected. From a single AFLP band only one sequence corresponded with the original AFLP marker. The other sequences were bands that shared not only the same size within 1-2 bp but also the same selective bases at the *Eco*RI and *Mse*I sites (100%).

15　　Further, some of the cloned sequences from within a band shared between 6 to 15 bp in common to each side (*Eco*RI and *Mse*I) of the original AFLP polymorphism (about 30% of bands).

　　　　To identify polymorphisms within the AFLP, the AFLP sequence was used to design primers to screen the Forrest *Bam*HI BAC library by PCR. For

20　　example, E$_{ATG}$M$_{CGA}$87 was a dominant AFLP band in coupling phase with the *rhg1* locus, and screening with a E$_{ATG}$M$_{CGA}$87 AFLP band primer yielded a single clone. Two internal primers were designed from the E$_{ATG}$M$_{CGA}$87 resistant allele and DNA from the corresponding BAC was used as template to extend the sequence from the AFLP marker both up and down stream by

sequencing. The sequence showed a single 5 bp indel underlay the polymorphic band and no SNPs were present. As used herein, an "indel" refers to a nucleotide insertion or a deletion (Figure 1B). No additional polymorphisms were detected in about 1,250 bp of flanking sequence.

5   Sequence comparison of both, resistant and the susceptible alleles of the co-dominant AFLP marker $E_{CTA}M_{AGG}113$ found polymorphisms including both indels and SNPs. There were 4 SNPs within 113 bp and 1 indel (21 bp) (Figure 1A). Primer sets were designed around the indel site and used to map the genetic position. The genetic position of the identified indel mapped to the

10   region of the original AFLP.

Sequence comparison of resistant and the susceptible alleles of the dominant AFLP marker $E_{CCC}M_{ATG}161$ found SNP polymorphism. There were 2 SNPs within 116 bp (Figure 1A). Primer sets were designed around the SNP site and used to map the genetic position. The genetic position of the identified

15   indel mapped to the region of the original AFLP.

Sequence comparison of both resistant and susceptible alleles of the dominant AFLP marker $E_{CCA}M_{AGC}114$ found SNP polymorphism adjacent to the *Eco*RI site. There was 1 SNP within 114 bp (Figure 1A).

Sequence comparison of resistant and susceptible alleles of the co-

20   dominant AFLP marker $E_{CCG}M_{AAC}405$ found polymorphisms including both indels and SNPs. There were 2 indels (12 bp and 4 bp) and 4 SNPs within 405 bp (Figure 1A). The 4 bp indel was two AG repeats in an $[AG]_5$ complex micro-satellite sequence. Primer sets were designed around both indel sites and used to map the genetic position. In both cases, the genetic position of the

identified indel mapped to the region of the original AFLP.

For the AFLP marker $E_{CGG}M_{AGA}116$, the polymorphisms were found adjacent to both the *Eco*RI and *Mse*I restriction sites (Figure 1A). The six selective nucleotide step was replaced by $M_{AGAGACT}$ and E_C. Using this primer

5   set the detection of the polymorphism on sequencing gels as well as the mapping of this sequence to the same location as the original AFLP was successful (Figure 2B). There was 1 indel (2 bp) and 1 SNPs within 116 bp (Figure 1A). The 2 bp indel was the $[A]_2$ extension of an $[A]_8$ repeat. Primer sets were designed around the indel and SNP sites and used to map their

10   genetic positions. In both cases, the genetic position of the identified polymorphism was identical to the region of the original AFLP.

Comparison of both alleles of the AFLP marker $E_{CCG}M_{AAC}405$ provided four SNPs, two indels and one SSR. The insertion of $[AG]_2$ in the $[AG]_8$ repeat of the resistance allele created a microsatellite polymorphism that was

15   designated SIUC-SAG405 by the present co-inventors. The difference of 4 bp between the two alleles at position 224 bp to 228 bp was enough to discriminate between the resistant and susceptible allele after electrophoresis through a 4% (v/w) Metaphor7 agarose gel. The 12 bp indel at 42 bp to 54 bp was used to design a sequence specific PCR marker (Figure 2D), and to

20   develop a TaqMan^{TM} assay for the *Rhg4* locus. SNPs were found within the $E_{CCG}M_{AAC}405$. The transversions of T at position 327 in the resistant allele to C at position 337 in the susceptible allele; and A at position 358 bp in the resistance allele to C at position 366 bp in the susceptible allele can also be used for high-throughput screening SNPs based assay.

An indel of 21 bp was responsible for the polymorphism at the $E_{CTA}M_{AGG}113$ AFLP locus between Essex and Forrest. PCR based markers were designed to flank the 21 bp indel and shown to be polymorphic, the new marker was named *CTA* (Figure 2C).

5   In the $E_{ATG}M_{CGA}87$ marker the insertion of CTTAT to form a tandem repeat in the Forrest allele at position 20 bp to 25 bp created a 5 bp polymorphism that was suitable for marker development. PCR primers were designed to develop a sequence specific PCR assay (Figure 2A), the new marker was named *ATG4*. The same indel was used to develop a TaqMan^{TM}

10   probe named *TMA5* to discriminate between the two alleles.

The genetic markers of the present invention can be used to reliably select SCN/SDS resistance, as described herein.

III.C.   SCN/SDS Resistance Genes

The term "gene" refers broadly to any segment of DNA associated with a

15   biological function. A gene encompasses sequences including but not limited to a coding sequence, a promoter region, a cis-regulatory sequence, a non-expressed DNA segment, a non-expressed DNA segment that contributes to gene expression, a DNA segment designed to have desired parameters, or combinations thereof. A gene can be obtained by a variety of methods,

20   including cloning from a biological sample, synthesis based on known or predicted sequence information, and recombinant derivation of an existing sequence.

The term "gene" thus includes an isolated soybean *rhg1* and SDS resistance gene as disclosed herein (Figure 3). The gene is capable of

conveying *Heterodera glycines*-infestation resistance or *Fusarium solani*-infection resistance to a non-resistant soybean germplasm, the gene located within a quantitative trait locus mapping to linkage group G and mapped by genetic markers of SEQ ID NOs:1-6, said gene located along said quantitative

5   trait locus between said markers. Positional cloning methods were used to isolate genomic sequences in the chromosomal regions of Forrest that confers SCN/SDS resistance, as further described in Example 4. Specifically, *rhg1* sequences were derived from BAC clones 21D9 and 73P6 of the Forrest *Bam*HI or *Hind*III BAC libraries (Meksem et al., 2000). Preferably, the gene

10   comprises the nucleotide sequence set forth as SEQ ID:13 (Figure 7A-B). BLASTP analysis of the conceptual translation of the *rhg1* gene (Figure 7C), set forth as SEQ ID:14 shows high homology to the <u>T46070</u> GenBank entry described as hypothetical protein T18N14.120 from *Arabidopsis thaliana* (Figure 7E-F), high homology to the rice Xa21 disease resistance gene

15   encoding a leucine-rich repeat protein, and high homology to the tomato CF-2 gene for resistance to *Cladosporium fulvus* (Figure 7D).

       The *rhg1* seqeunces disclosed herein can also be used to isolate *rhg1* cDNAs according to methods well-known in the art. A representative *rhg1* partial cDNA is set forth as SEQ ID NO:122. This segment of the *rhg1* gene

20   shows homology to the leucine-rich regions of the *Arabidopsis* hypothetical protein T18N14.120 (Gen Bank T46070) and tomato CF-2 resistance genes.

       For example, the term "gene" also includes an isolated soybean *Rhg4* gene. The gene is capable of conveying *Heterodera glycines*-infestation resistance to a non-resistant soybean germplasm, said gene located within a

quantitative trait locus mapping to linkage group A2 and mapped by the AFLP markers of SEQ ID NOs:6-12, said gene located along said quantitative trait locus between said markers. Preferably, the gene comprises a nucleotide sequence set forth as any one of SEQ ID NOs:16-19.

5   Genes underlying quantitative traits, or genes with related function, such as disease resistance, are often organized in clusters within the genome (e.g., Staskawicz (1995) *Science* 268:661-667). In the case of SCN/SDS resistance, previous studies by the co-inventors of the present invention have suggested that the resistance trait in Forrest may be caused by four genes in a cluster with

10   two pairs in close linkage or by a two-gene cluster with each gene displaying pleitropy (Meksem et al., 1999). Thus, genomic DNA isolated and disclosed herein comprise multiple resistance gene sequences. Additional sequences derived from the SCN/SDS resistance locus are set forth as SEQ ID NOs:20-66. BLASTX analysis of these sequences reveals further homology to known

15   proteins in other organisms, supporting that they comprise new partial gene sequences (Table 1). Of particular interest, BLASTX analysis of the sequences set forth as SEQ ID NOs:67-114 reveals that several of the disclosed sequences have high homology to the T46070 GenBank entry described as hypothetical protein T18N14.120 from *Arabidopsis thaliana*, high homology to

20   the tomato CF-2 disease resistance genes encoding leucine-rich repeat proteins, and to the tomato CF-9 gene for resistance to *Cladosporium fulvus* (Table 1).

The present invention also pertains to resistance genes related to *rhg1* and *Rhg4*. Partial cDNAs of additional putative SCN/SDS resistance genes, set

forth as SEQ ID NOs:67-114, were identified based on hybridization to *rhg1* and

Rhg4 sequences, as further described in Example 5. BLASTX analysis of

these sequences reveals further homology to known proteins in other

organisms, supporting that they comprise new partial gene sequences (Table

5   2). Of particular interest, BLASTX analysis of the sequences set forth as SEQ

ID NOs:67-114 reveals that several of the disclosed sequences have high

homology to the T46070 GenBank entry described as hypothetical protein

T18N14.120 from *Arabidopsis thaliana*, high homology to the tomato CF-2

disease resistance genes encoding leucine-rich repeat proteins, and to the

10   tomato CF-9 gene for resistance to *Cladosporium fulvus* (Table 2). Based on

their hybridization to *rhg1* and *Rhg4* sequences, genes comprising any of SEQ

ID NOs:67-114 may also confer resistance to race 3 *Heterodera glycines*. It will

be apparent to one having ordinary skill in the art that the disclosed sequences,

or portion thereof, can be used to identify, confirm and/or screen for SDS, SCN

15   and/or other resistance or for loci that confer SDS, SCN and/or other

resistance.

Table 1

SEQ ID NO.	inventor's reference	best BLAST hit (ACCESSION)	Score (bits)	E value	Identitties	Positives
20	III-00_F2-3RCF1900-2450	T47727	230	9e-60	114/170 (67%)	134/170 (78%)
21	III-01_21d9A1,1A1	no significant similarity				
22	III-01_21d9A2,11F11Rlaccase	AC007063	97	1e-19	62/166 (37%)	92/166 (55%)
23	III-01_21d9A2,4A4Mic	no significant similarity				
24	III-01_CMG,smalF1-1F	T46070	67	4e-13	49/147 (33%)	62/147 (41%)
25	III-02_21d9A2,12A12FNaH+hypoth	T00576	67	2e-10	57/188 (30%)	87/188 (45%)
26	III-02_F3-1RCF2000-2500	T46070	170	7e-42	79/105 (75%)	93/105 (88%)
27	III-03_21d9A1,1E1Flaccase	AC007020	61	1e-08	37/65(56%)	43/65 (65%)
28	III-03_21d9A2,12A12RNaH+hypothet	AC007063	116	2e-25	61/165 (36%)	95/165 (56%)
29	III-03_21d9A2,4B4ESTM	no significant similarity				
30	III-03_21d9A2,8F8CF1a	T47727	187	53-48	95/142 (66%)	106/142 (73%)
31	III-03_21d9A2,8F8CFHomol	T47727	177	5e-45	90/132 (68%)	100/132 (75%)
32	III-03_CMG,smalF1-3FCF300-1100	T46070	107	4e-27	67/189 (35%)	89/189 (46%)
33	III-03_F3-2R1800-Cterm	T47727	201	1e-64	97/129 (75%)	113/129 (87%)
34	III-04_21d9A1,1E1R	no significant similarity				
35	III-04_21d9A2,1B1	no significant similarity				
36	III-04_21d9A2,6D6mic	no significant similarity				
37	III-05_21d9A1,1C1GmxLaccase	AB010692	153	2e-36	80/124 (64%)	90/124 (72%)
38	III-05_21d9A2,4C4CFHomol	T46070	125	6e-28	65/106 (61%)	72/106 (67%)
39	III-06_21d9A2,11A11laccasegene	AC007020	67	3e-12	30/49 (61%)	35/49 (71%)
40	III-07_21d9A1,2A2F	no significant similarity				
41	III-08_21d9A1,2A2R	no significant similarity				

	Name	Hit	Score	E-value	Identities	Positives
42	III-08_21d9A2,6F6	no significant similarity				
43	III-09_21d9A1,1E1	no significant similarity				
44	III-09_21d9A1,2D2FNaH+hypothe	AC007063	84	93-17	44/127 (34%)	74/127 (57%)
45	III-09_21d9A2,4E4Laccase	AC007020	90	1e-32	43/53 (81%)	46/53 (86%)
46	III-09_21d9A2,9A9	no significant similarity				
47	III-10_21d9A2,11C11	T47325	53	3e-06	45/132 (34%)	65/132 (49%)
48	III-10_21d9A2,11C11hypothetical	T47325	53	3e-06	45/132 (34%)	65/132 (49%)
49	III-11_21d9A1,1F1SatAT	no significant similarity				
50	III-11_21d9A2,4A4F	no significant similarity				
51	III-11_21d9A2,4F4SatTA	no significant similarity				
52	III-12_21d9A2,1F1NaHexchangine	AC007063	126	3e-28	72/181 (39%)	108/181(58%)
53	III-12_21d9A2,4A4RSatTAGA	no significant similarity				
54	III-13_21d9A1,1G1NaHexchanHypothe	T00576	50	2e-05	31/83 (37%)	44/83 (52%)
55	III-13_21d9A1,8D8CF500-1000	T46070	84	4e-24	48/127 (37%)	66/127 (51%)
56	III-13_21d9A2,4B4FSatGAAAA	no significant similarity				
57	III-14_21d9A2,11E11GmxEST	no significant similarity				
58	III-14_21d9A2,1G1	no significant similarity				
59	III-15_21d9A1,8E8	no significant similarity				
60	III-15_21d9A2,4C4FCF1600-1000	T46070	158	6e-38	99/215 (46%)	113/215 (52%)
61	III-15_21d9A2,9D9NaHIonexch	AC007063	64	1e-09	38/118 (32%)	59/118 (49%)
62	III-16_21d9A1,11D11laccase	CAA74104	82	4e-17	35/49 (71%)	43/49 (87%)
63	III-16_21d9A2,11F11MicSatTA	no significant similarity				
64	III-16_21d9A2,4C4R300-1000	T46070	110	3e-32	67/178 (37%)	86/178 (47%)
65	III-17_21d9A1,2A2SatGA	no significant similarity				
66	III-17_21d9A1,2A2SatTAA	no significant similarity				
73	II-01F2-4RCf1900-2400	T46070	187	6e-47	99/183 (54%)	123/183 (67%)

Table 2

SEQ ID NO.	inventor's reference	best BLAST hit (ACCESSION)	Score (bits)	E value	Identities	Positives
67	3A Cf2 homologues to the +2ORF clone ID:07d9	T47727	189	4e-47	103/215 (47%)	127/215 (58%)
68	3B Cf2 homologues to the -2ORF clone ID.05d7	T46070	148	8e-35	76/157 (48%)	98/157 (62%)
69	3C Cf2 homologues to the +3 ORF clone ID:17P9	T47727	200	2e-50	100/136 (73%)	113/136 (82%)
70	3D Cf2 homologues to the -3ORF clone ID:06d8	T46070	163	2e-39	86/179 (48%)	110/179 (61%)
71	II-00_F2-3RCF1900-2450	T47727	230	9e-60	114/170 (67%)	134/170 (78%)
72	II-01CMGsmaIF1-1F300-1000	T46070	76	4e-13	49/147 (33%)	62/147 (41%)
73	II-01F2-4RCf1900-2400	T46070	187	6e-47	99/183 (54%)	123/183 (67%)
74	II-02F3-1RCF2000-2500	T46070	170	7e-42	79/105 (75%)	93/105 (88%)
75	II-03.21dA2,8F8CF1-500	T47727	187	5e-48	95/142 (66%)	106/142 (73%)
76	II-03CMG,smaIF1-3FCF300-1100	T46070	107	4e-27	67/189 (35%)	89/189 (46%)
77	II-03F3-2R1800-Cterm	T47727	201	1e-64	97/129 (75%)	113/129 (87%)
78	II-04.21dA1,1E1R	no significant similarity				
79	II-05.21dA2,4C4CFhomol	T46070	125	6e-28	65/106 (61%)	72/106 (67%)
80	II-12CFLNO1F-CFNO1F	T46070	135	2e-33	74/165 (44%)	97/165 (57%)
81	II-12CFLNO1F-CFLNOIR	T46070	273	2e-72	133/183 (72%)	156/183 (84%)
82	II-12CFLNO1F-CFLNNIF	T46070	184	7e-46	91/128 (71%)	100/128 (78%)
83	II-12CFLNO1F-CFLNN2F	T46070	109	3e-24	69/189 (36%)	89/189 (46%)
84	II-13.21dA1,8D8CF500-1000	T46070	84	4e-24	48/127 (37%)	66/127 (51%)
85	II-15.21dA2,4C4FCF1600-1000	T46070	158	6e-38	99/215 (46%)	113/215 (52%)
86	II-29.21dA2,8F8FCF500upstream	T47727	102	2e-39	56/105 (53%)	67/105 (63%)
87	II-30.21d9A2,12E12ESTMedicago	T47731	238	6e-62	119/163 (73%)	132/163 (80%)
88	II-30.21d9A2,8F8RCFpromoter	no significant similarity				
89	II-30.E2,TetRP1downstreamtoRhg1	S05434	35	1.0	30/109 (27%)	49/109 (44%)
90	II-32.E3,TetRP1CF1115-1249	no significant similarity				

91	II-Cf homol-01CMGsmalF1-2F	T46070	76	4e-13	49/147 (33%)	62/147 (41%)
92	II-Cf homol-CMGsmalF1-2F	T46070	125	8e-32	74/188 (39%)	95/188 (50%)
93	II-Cf homol-03CMGsmalF1-3	T46070	105	1e-26	66/188 (35%)	88/188 (46%)
94	II-Cf homol-06CMGsmalF2-2F	T46070	123	2e-27	80/224 (35%)	105/224 (46%)
95	II-Cf homol-07CMGsmalF2-3F	T46070	123	2e-27	80/224 (35%)	105/224 (46%)
96	II-Cf homol-08CMGsmalF2-4F03	T46070	118	6e-29	71/183 (38%)	90/183 (48%)
97	II-Cf homol-10CMGsmalF3-2F	T46070	184	7e-46	91/128 (71%)	100/128 (78%)
98	II-Cf homol-09CMGsmalF3-1F	T46070	184	6e-46	91/128 (71%)	100/128 (78%)
99	II-Cf homol-smalF3-3F	T46070	265	2e-70	128/174 (73%)	151/174 (86%)
100	II-Cf homol-12CMGsmalF3-4F	T46070	184	7e-46	89/107 (83%)	97/107 (90%)
101	II-Cf homol-13CMGsmalF1-1R	T46070	279	3e-74	136/191 (71%)	159/191 (83%)
102	II-Cf homol-14CMGsmalF1-2R	T46070	261	3e-69	127/176 (72%)	148/176 (83%)
103	II-Cf homol-15CMGsmalF1-3R	T47727	246	1e-64	120/162 (74%)	140/162 (86%)
104	II-Cf homol-16CMGsmalF1-4R	T46070	263	1e-70	128/176 (72%)	149/176 (83%)
105	II-Cf homol-17CMGsmalF2-1R	T46070	268	5e-71	131/183 (71%)	155/183 (84%)
106	II-Cf homol-18CMGsmalF2-2R	T46070	244	4e-65	118/159 (74%)	137/159 (85%)
107	II-Cf homol-05F3-4R	T46070	187	6e-47	90/136 (66%)	111/136 (81%)
108	II-Cf homol-00F2-3R	T46070	224	3e-58	108/148 (72%)	127/148 (84%)
109	II-Cf homol-01F2-4R	T46070	187	6e-47	99/183 (54%)	123/183 (67%)
110	II-Cf homol-02F3-1R	T46070	170	7e-42	79/105 (75%)	93/105 (88%)
111	II-Cf homol-03F3-2R	T47727	202	9e-65	97/133 (72%)	11/133 (84%)
114	II-Cf homol-04F3-3R	T46070	128	1e-30	65/108 (60%)	72/108 (66%)
114	II-Cf homol-05CMGsmalF2-F	T46070	184	6e-46	91/128 (71%)	100/128 (78%)
114	II-downstream to Rhg1	no significant similarity				

III.D.   SCN/SDS Resistance Gene Promoters

The term "promoter region" defines a nucleotide sequence within a gene that is positioned 5' to a coding sequence of a same gene and functions to direct transcription of the coding sequence. The promoter region includes a transcriptional start site and at least one cis-regulatory element. The present invention encompasses nucleic acid sequences that comprise a promoter region of an SCN/SDS resistance gene, or functional portion thereof.

The terms "cis-acting regulatory sequence" or "cis-regulatory motif" or "response element", as used herein, each refer to a nucleotide sequence that enables responsiveness to a regulatory transcription factor. Responsiveness can encompass a decrease or an increase in transcriptional output and is mediated by binding of the transcription factor to the DNA molecule comprising the response element.

The term "transcription factor" generally refers to a protein that modulates gene expression by interaction with the cis-regulatory element and cellular components for transcription, including RNA Polymerase, Transcription Associated Factors (TAFs), chromatin-remodeling proteins, and any other relevant protein that impacts gene transcription.

The term "gene expression" generally refers to the cellular processes by which a biologically active polypeptide is produced from a DNA sequence.

A "functional portion" of a promoter gene fragment is a nucleotide sequence within a promoter region that is required for normal gene transcription. To determine nucleotide sequences that are functional, the

expression of a reporter gene is assayed when variably placed under the direction of a promoter region fragment.

Promoter region fragments can be conveniently made by enzymatic digestion of a larger fragment using restriction endonucleases or DNAse I.

5   Preferably, a functional promoter region fragment comprises about 5,000 nucleotides, more preferably 2,000 nucleotides, more preferably about 1,000 nucleotides, more preferably a functional promoter region fragment comprises about 500 nucleotides, even more preferably a functional promoter region fragment comprises about 100 nucleotides, and even more preferably a

10   functional promoter region fragment comprises about 20 nucleotides.

Within a candidate promoter region or response element, the presence of regulatory proteins bound to a nucleic acid sequence can be detected using a variety of methods well known to those skilled in the art (Ausubel et al., 1992). Briefly, *in vivo* footprinting assays demonstrate protection of DNA sequences

15   from chemical and enzymatic modification within living or permeabilized cells. Similarly, *in vitro* footprinting assays show protection of DNA sequences from chemical or enzymatic modification using protein extracts. Nitrocellulose filter-binding assays and gel electrophoresis mobility shift assays (EMSAs) track the presence of radiolabeled regulatory DNA elements based on provision of

20   candidate transcription factors.

The terms "reporter gene" or "marker gene" or "selectable marker" each refer to a heterologous gene encoding a product that is readily observed and/or quantitated. A reporter gene is heterologous in that it originates from a source foreign to an intended host cell or, if from the same source, is modified from its

original form. Non-limiting examples of detectable reporter genes that can be operably linked to a transcriptional regulatory region can be found in brown and PCT International Publication No. WO 97/47763. Preferred reporter genes for transcriptional analyses include the lacZ gene (See, e.g., Rose & Botstein

5   (1983) *Meth Enzymol* 101:167-180), Green Fluorescent Protein (GFP) (Cubitt et al. (1995) *Trends Biochem Sci* 20:448-455), luciferase, or chloramphenicol acetyl transferase (CAT). Preferred reporter genes for stable transformation include but are not limited to antibiotic resistance genes. Any suitable reporter and detection method can be used, and it will be appreciated by one of skill in

10   the art that no particular choice is essential to or a limitation of the present invention.

An amount of reporter gene can be assayed by any method for qualitatively or preferably, quantitatively determining presence or activity of the reporter gene product. The amount of reporter gene expression directed by

15   each test promoter region fragment is compared to an amount of reporter gene expression to a control construct comprising the reporter gene in the absence of a promoter region fragment. A promoter region fragment is identified as having promoter activity when there is significant increase in an amount of reporter gene expression in a test construct as compared to a control construct.

20   The term "significant increase", as used herein, refers to an quantified change in a measurable quality that is larger than the margin of error inherent in the measurement technique, preferably an increase by about 2-fold or greater relative to a control measurement, more preferably an increase by about 5-fold or greater, and most preferably an increase by about 10-fold or greater.

A representative SCN/SDS resistance gene promoter, the *rhg1* promoter, is set forth as SEQ ID NO:15. The *rhg1* promoter is useful for directing gene expression of heterologous sequences *in vivo* or in assays to identify modulators of *rhg1* expression, described further herein below.

5   The present invention further provides an isolated SCN/SDS resistance gene promoter region, or functional portion thereof, comprising an about 90 kb fragment of soybean genomic clone 73P6 between *Bam*HI restriction sites and 21d9 between *Hin*DIII restriction site. The genomic clone is available from the Forrest BAC library described in Meksem et al (2000), *Theor Appl Genet* 101

10   5/6: 747-755, available through Southern Illinois University-Carbondale (Carbondale, Illinois), Texas A&M University BAC center (College Station, Texas), and Research Genetics (Huntsville, Alabama). An isolated SCN/SDS resistance gene promoter region, or functional portion thereof, comprising an about 4.5 kb fragment of soybean genomic clone 21d9A2 8F8 between EcoRI

15   restriction sites is also disclosed.

III.E.   Chimeric Genes

The present invention also encompasses chimeric genes comprising the disclosed SCN/SDS resistance gene sequences. The term "chimeric gene", as used herein, refers to an SCN/SDS resistance gene promoter region operably

20   linked to an open reading frame, wherein the nucleotide sequence created is not naturally occurring. In this regard, the open reading frame is also described as a "heterologous sequence". The term "chimeric gene" also encompasses a promoter region operably linked to an SCN/SDS resistance gene coding sequence, a nucleotide sequence producing an antisense RNA molecule, a

RNA molecule having tertiary structure, such as a hairpin structure, or a double-stranded RNA molecule.

The term "operably linked", as used herein, refers to a promoter region that is connected to a nucleotide sequence in such a way that the transcription of that nucleotide sequence is controlled and regulated by that promoter region. Techniques for operatively linking a promoter region to a nucleotide sequence are well known in the art.

The terms "heterologous gene", "heterologous DNA sequence", "heterologous nucleotide sequence", "exogenous nucleic acid molecule", or "exogenous DNA segment", as used herein, each refer to a sequence that originates from a source foreign to an intended host cell or, if from the same source, is modified from its original form. Thus, a heterologous gene in a host cell includes a gene that is endogenous to the particular host cell but has been modified, for example by mutagenesis or by isolation from native cis-regulatory sequences. The terms also include non-naturally occurring multiple copies of a naturally occurring nucleotide sequence. Thus, the terms refer to a DNA segment that is foreign or heterologous to the cell, or homologous to the cell but in a position within the host cell nucleic acid wherein the element is not ordinarily found.

IV. Polypeptide Sequences of SCN/SDS Resistance Proteins

The polypeptides provided by the present invention include the isolated polypeptide of SEQ ID NO:14, fusion proteins comprising SCN/SDS resistance gene amino acid sequences, biologically functional analogs, and polypeptides

that cross-react with an antibody that specifically recognizes an SCN/SDS resistance gene polypeptide.

The term "isolated", as used in the context of a polypeptide, indicates that the polypeptide exists apart from its native environment and is not a

5 product of nature. An isolated polypeptide can exist in a purified form or can exist in a non-native environment such as, for example, in a transgenic host cell.

The term "purified", when applied to a polypeptide, denotes that the polypeptide is essentially free of other cellular components with which it is

10 associated in the natural state. Preferably, a polypeptide is a homogeneous solid or aqueous solution. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A polypeptide that is the predominant species present in a preparation is substantially purified.

15 The term "purified" denotes that a polypeptide gives rise to essentially one band in an electrophoretic gel. Particularly, it means that the polypeptide is at least about 50% pure, more preferably at least about 85% pure, and most preferably at least about 99% pure.

The term "substantially identical" in the context of two or more

20 polypeptides sequences is measured by (a) polypeptide sequences having about 35%, or 45%, or preferably from 45-55%, or more preferably 55-65%, or most preferably 65% or greater amino acids that are identical or functionally equivalent. Percent "identity" and methods for determining identity are defined herein under the heading Nucleotide and Amino Acid Sequence Comparisons.

Substantially identical polypeptides also encompass two or more polypeptides sharing a conserved three-dimensional structure. Computational methods can be used to compare structural representations, and structural superpositions can be generated and easily tuned to identify similarities around

5   important active sites or ligand binding sites. See Henikoff et al. (2000) *Electrophoresis* 21(9):1700-1706; Huang et al. (2000) *Pac Symp Biocomput* 230-241; Saqi et al.,1999; and Barton (1998) *Acta Crystallogr D Biol Crystallogr* 54:1139-1146.

The term "functionally equivalent" in the context of amino acid

10   sequences is well known in the art and is based on the relative similarity of the amino acid side-chain substituents. See Henikoff and Henikoff (2000) *Adv Protein Chem* 54:73-97. Relevant factors for consideration include side-chain hydrophobicity, hydrophilicity, charge, and size. For example, arginine, lysine, and histidine are all positively charged residues; that alanine, glycine, and

15   serine are all of similar size; and that phenylalanine, tryptophan, and tyrosine all have a generally similar shape. By this analysis, described further herein below, arginine, lysine, and histidine; alanine, glycine, and serine; and phenylalanine, tryptophan, and tyrosine; are defined herein as biologically functional equivalents.

20   In making biologically functional equivalent amino acid substitutions, the hydropathic index of amino acids can be considered. Each amino acid has been assigned a hydropathic index on the basis of their hydrophobicity and charge characteristics, these are: isoleucine (+ 4.5); valine (+ 4.2); leucine (+ 3.8); phenylalanine (+ 2.8); cysteine (+ 2.5); methionine (+ 1.9); alanine (+ 1.8);

glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

The importance of the hydropathic amino acid index in conferring interactive biological function on a protein is generally understood in the art (Kyte et al. (1982) *J Mol Biol* 157:105.). It is known that certain amino acids can be substituted for other amino acids having a similar hydropathic index or score and still retain a similar biological activity. In making changes based upon the hydropathic index, the substitution of amino acids whose hydropathic indices are within ±2 of the original value is preferred, those which are within ±1 of the original value are particularly preferred, and those within ±0.5 of the original value are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Patent No. 4,554,101 states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with its immunogenicity and antigenicity, *i.e.* with a biological property of the protein. It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent protein.

As detailed in U.S. Patent No. 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+ 3.0); lysine (+ 3.0); aspartate (+ 3.0±1); glutamate (+ 3.0±1); serine (+ 0.3); asparagine (+ 0.2); glutamine (+ 0.2); glycine (0); threonine (-0.4); proline (-0.5±1); alanine (-

0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4).

In making changes based upon similar hydrophilicity values, the substitution of amino acids whose hydrophilicity values are within ±2 of the original value is preferred, those which are within ±1 of the original value are particularly preferred, and those within ±0.5 of the original value are even more particularly preferred.

The present invention also encompasses SCN/SDS resistance gene polypeptide fragments or functional portions of an SCN/SDS resistance gene polypeptide. Such functional portion need not comprise all or substantially all of the amino acid sequence of a native resistance gene product. The term "functional" includes any biological activity or feature of SCN/SDS resistance gene, including immunogenicity.

The present invention also includes longer sequences comprising an SCN/SDS resistance gene polypeptide, or portion thereof. For example, one or more amino acids can be added to the N-terminal or C-terminal of an SCN/SDS resistance gene polypeptide. Fusion proteins comprising SCN/SDS resistance gene polypeptide sequences are also provided within the scope of the present invention. Methods of preparing such proteins are known in the art.

The present invention also encompasses functional analogs of an SCN/SDS resistance gene polypeptide. Functional analogs share at least one biological function with an SCN/SDS resistance gene polypeptide. An exemplary function is immunogenicity. In the context of amino acid sequence, biologically functional analogs, as used herein, are peptides in which certain,

but not most or all, of the amino acids can be substituted. Functional analogs

can be created at the level of the corresponding nucleic acid molecule, altering

such sequence to encode desired amino acid changes. In one embodiment,

changes can be introduced to improve the antigenicity of the protein. In

5   another embodiment, an SCN/SDS resistance gene polypeptide sequence is

varied so as to assess the activity of a mutant SCN/SDS resistance gene

polypeptide. In still another embodiment, amino acid changes can be made to

improve the stability of the polypeptide.

Isolated polypeptides and recombinantly produced polypeptides can be

10   purified and characterized using a variety of standard techniques that are well

known to the skilled artisan. See, e.g. Ausubel et al. (1992); Bodanszky et al.,

1976; and Zimmer et al. (1993) Peptides, pp. 393B394, ESCOM Science

Publishers, B. V.

V. Nucleotide and Amino Acid Sequence Comparisons

15   The terms "identical" or percent "identity" in the context of two or more

nucleotide or polypeptide sequences, refer to two or more sequences or

subsequences that are the same or have a specified percentage of amino acid

residues or nucleotides that are the same, when compared and aligned for

maximum correspondence, as measured using one of the sequence

20   comparison algorithms disclosed herein or by visual inspection.

The term "substantially identical" in regards to a nucleotide or

polypeptide sequence means that a particular sequence varies from the

sequence of a naturally occurring sequence by one or more deletions,

substitutions, or additions, the net effect of which is to retain at least some of

biological activity of the natural gene, gene product, or sequence. Such sequences include "mutant" sequences, or sequences wherein the biological activity is altered to some degree but retains at least some of the original biological activity. The term "naturally occurring", as used herein, is used to

5   describe a composition that can be found in nature as distinct from being artificially produced by man. For example, a protein or nucleotide sequence present in an organism, which can be isolated from a source in nature and which has not been intentionally modified by man in the laboratory, is naturally occurring.

10   For sequence comparison, typically one sequence is regarded as a reference sequence to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer program, subsequence coordinates are designated if necessary, and sequence algorithm program parameters are selected. The sequence

15   comparison algorithm then calculates the percent sequence identity for the designated test sequence(s) relative to the reference sequence, based on the selected program parameters.

Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman (1981) *Adv Appl Math*

20   2:482, by the homology alignment algorithm of Needleman & Wunsch (1970) *J Mol Biol* 48:443, by the search for similarity method of Pearson & Lipman (1988) *Proc Natl Acad Sci USA* 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin

Genetics Software Package, Genetics Computer Group, Madison, WI), or by visual inspection. See generally, Ausubel et al. (1992).

A preferred algorithm for determining percent sequence identity and sequence similarity is the BLAST algorithm, which is described in Altschul et al.

5   (1990) J Mol Biol 215: 403-410. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold

10   score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold. These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased.

15   Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when the cumulative

20   alignment score falls off by the quantity X from its maximum achieved value, the cumulative score goes to zero or below due to the accumulation of one or more negative-scoring residue alignments, or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses

as defaults a wordlength W=11, an expectation E=10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix. See Henikoff and Henikoff (1989) *Proc*

5   *Natl Acad Sci USA* 89:10915.

In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences. See, *e.g.,* Karlin and Altschul (1993) *Proc Natl Acad Sci USA* 90:5873-5887. One measure of similarity provided by the BLAST algorithm is the smallest sum

10   probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a test nucleic acid sequence is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid sequence to the reference nucleic acid sequence is less than about

15   0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

VI.   Method for Detecting a Nucleic Acid Molecule Associated with SCN/SDS Resistance

In another aspect of the invention, a method is provided for detecting a

20   nucleic acid molecule that encodes an SCN/SDS resistance polypeptide. Such methods can be used to detect SCN/SDS resistance gene variants and related resistance gene sequences. The disclosed methods facilitate genotyping, cloning, gene mapping, and gene expression studies.

VI.A. Genetic Variants

In one embodiment, genetic assays based on nucleic acid molecules of the present invention can be used to screen for genetic variants by a number of PCR-based techniques, including single-strand conformation polymorphism

5   (SSCP) analysis (Orita et al. (1989) *Proc Natl Acad Sci USA* 86(8):2766-2770), SSCP/heteroduplex analysis, enzyme mismatch cleavage, direct sequence analysis of amplified exons (Kestila et al. (1998) *Mol Cell* 1(4):575-582; Yuan et al. (1999) *Hum Mutat* 14(5):440-446), allele-specific hybridization (Stoneking et al. (1991) *Am J Hum Genet* 48(2):370-82), and restriction analysis of amplified

10   genomic DNA containing the specific mutation. Automated methods can also be applied to large-scale characterization of single nucleotide polymorphisms (Brookes (1999) *Gene* 234(2):177-186; Wang et al. (1998) *Science* 280(5366):1077-82). Preferred detection methods are non-electrophoretic, including, for example, the TaqMan™ allelic discrimination assay, PCR-OLA,

15   molecular beacons, padlock probes, and well fluorescence. See Landegren et al. (1998) *Genome Res* 8:769-776.

In a preferred embodiment, genetic markers for SCN/SDS resistance disclosed herein are used in a PCR-based genotyping assay, preferably, a TaqMan™ assay as disclosed in Example 6. The TaqMan™ allelic

20   discrimination assay is based on the 5' nuclease activity of Taq polymerase and detection of a fluorescent reporter during or after PCR reactions (Livak et al. (1995) *PCR Meth and Applic* 4:357-362; Livak et al. (1995) *Nat Genet* 9:341-342). Each TaqMan™ probe consists of a 25-35 base oligonucleotide complementary to one of two alleles with a 3' quencher dye attached (6-

carboxy-N,N,N'5N' tetrachlorofluorescein; TAMRA). The oligomer complimentary to allele 1 is linked covalently to a 5' reporter dye (6-carboxy-4,7,2',7', tetrachlorofluorescenin; TET) while allele 2 is linked to a dye that fluoresces at a distinct wavelength (6-carboxyfluorescein; FAM). PCR directed

5   by flanking oligomers of 18-20 bases causes degradation during the extension phase of the oligomer that hybridizes most efficiently to the polymorphic site(s) in the sample. Adaptations can make the assay chemistry suitable for multiplexing (Nasarabadi et al. (1999) *BioTechniques* 27:1116-1117) and miniaturization (Kalinina et al. (1997) *Nucl Acids Res* 25:1999-2004) to reduce

10   cost and increase throughput.

The present invention discloses sequences suitable for use with the TaqMan[TM] method for genotyping SCN/SDS resistance, further disclosed in Example 6. As one example, the TaqMan[TM] assay was used to distinguish between two insertion polymorphisms in alleles of an AFLP marker that is

15   located about 50 kbp from the *Rhg4* gene (Figure 4). Genomic DNA samples were analyzed using the TaqMan[TM] PCR protocol (Livak et al., 1995a, 1995b). Using the raw fluorescence signals of the reporter dyes FAM and TET from the "dye component" field of the sequence detection software, two grouping methods were performed. Each method detected four distinct populations

20   (Figure 5). The four populations could be assigned according to the FAM:TET ratio based on where the heterogeneous class cut-off was placed.

For the TaqMan[TM] selection, two grouping methods were arbitrarily selected to attempt to accurately separate heterogeneous lines from homogeneous lines at each allele. For grouping method 1 (Taqman[TM] 1) a

stringent cut-off was used to reduce the number called as potentially heterogeneous. Fluorophore ratios were as follows; no amplification (FAM and TET both less than 6 units); allele 1 homozygous (FAM less than 7, TET greater than 7); allele 2 homozygous (FAM greater than 10, TET less than 5);

5 and heterogeneous for allele 1 and allele 2 (FAM greater than 7, TET 5-8). For TaqMan™ selection grouping method 2 (TaqMan™ 2), a lower stringency cut-off value was used to increase the number called as potentially heterogeneous. Ratios were: no amplification (FAM and TET both less than 6 units); allele 1 homozygous (FAM less than 5, TET greater than 7); allele 2 homozygous (FAM

10 greater than10, TET less than 5); and heterogeneous for allele 1 and allele 2 (FAM greater than 5, TET 5-9).

Based on the FI of the ExF RIL population, the 86 selected individuals were classified into 3 classes: 15 resistant, 60 susceptible and 11 segregating lines. TaqMan™ analysis of 86 individuals from the RILs by method 1 (high

15 stringency) shows a strong agreement between allele 1 and susceptibility to SCN (56 from the 60 susceptible lines were allele 1 type). However, there was lesser agreement between allele 2 and resistance to SCN (only 15 lines from the 23 lines showing the presence of allele 2 were resistant by phenotype) due to the segregation of *rhg1*, the second gene necessary for resistance to SCN in

20 Forrest. Of the 11 lines known to be heterogeneous for the resistance to SCN phenotype, five should segregate at *Rhg4*. TaqManTM method 1 identified one among the five classified as heterogenous (the 5 include 4 miss-classified lines, see below). TaqManTM method 2 identified all five among the 11 classified as heterogenous, however the 11 include 6 miss-classified lines.

To validate the specificity of TaqMan[TM] genotyping, samples of each of the RILs classified by the TaqMan[TM] method (Figure 5) were re-scored by PCR and gel electrophoresis (Figure 6) according to methods described in Example 7. The classifications produced by the two methods agreed with Taqman[TM] assay 1 most closely but with eight exceptions. The miss-scores were as follows (annotated as RIL#; FI phenotype; allele with TaqManTM grouping method 2; allele with TaqMan[TM] grouping method 1; allele by gel marker score): 4;S;H;H;S: 21;R;H;H;R: 32;R;H;H;R: 44;S;S;S;H: 51;S;S;S;H: 59;R;H;H;R: 63;S;S;S;R: 78;R;H;H;R.

The majority of disagreements resulted from resistant lines that were scored as heterogeneous by TaqMan[TM] but not gel electrophoresis or phenotype (4 of 8) and phenotypically susceptible lines that were scored incorrectly by gel electrophoresis (3 of 8). One genotype (RIL84) was miss-scored relative to phenotype (84SRRR) by all the allele genotyping methods and may represent a recombination event between A2D8 and *Rhg4*.

The genoytpe and phenotype were generally in close agreement among the eighty six genomic DNA samples analyzed using the TaqMan[TM] PCR protocol. The lesser agreement between Allele 2 and resistance to SCN (15 of 23) was shown to be due to the segregation of *rhg1*, by scoring of the BARC-Satt 309 marker (Meksem et al., 1999). The bias toward a higher frequency of allele 1 is caused by sampling error (Chang et al., 1997). The accuracy of genotyping was high by the TaqMan[TM] assay and was better than one pass gel electrophoresis (Prabhu et al., 1999). Even compared to a highly optimized gel electrophoresis assay reported herein the assays were not significantly different

in accuracy for detecting the genotypes within the F_5 derived RILs in a single pass assay. Exactly 78 of the 86 tested with both, TaqMan[TM] and gel electrophoresis results agreed. There were 5 errors with Taqman[TM] (94% accurate) and 3 errors with gel electrophoresis (96% accurate) judged by

5   replicated genotyping (not shown) and the phenotype. Low frequencies of error are important to the accurate selection of resistance (Cregan et al., 1999a; Prabhu et al., 1999) and in the generation of accurate genetic maps (Cregan et al., 1999b).

VI.B.   Cloning of SCN/SDS Resistance Genes and Related Genes

10   The nucleic acids of the present invention can be used to clone genes and genomic DNA comprising the sequences. Alternatively, the nucleic acids of the present invention can be used to clone genes and genomic DNA of related sequences. For this purpose, representative probes, hybridization conditions, and PCR primers are described in the section entitled Nucleotide

15   Sequences of SCN/SDS Resistance Genes and Associated Markers herein above and in Examples 4 and 5. Preferably, the nucleic acids used for this method comprise sequences set forth as any one of SEQ ID NOs:13, 15-114, more preferably SEQ ID NOs: 13 and 16-19.

In another embodiment, the present invention provides a method of

20   positional cloning of genes and other sequences located adjacent or near the disclosed sequences within the soybean genome. The method comprises: (a) identifying a first nucleic acid genetically linked to a SCN/SDS resistance locus; and (b) cloning the first nucleic acid. Optionally, the first nucleic acid can comprise the rhg1 and SDS locus or the Rhg4 locus. Preferably, the SCN/SDS

resistance locus corresponds to a nucleic acid selected from any one of SEQ ID NOs:13 and 16-19.

Positional cloning first involves creating a physical map of a contig (contiguous overlapping of cloned DNA inserts), in the genomic region encompassing one or more marker loci and the target gene. The target gene is then identified and isolated within one or more clones residing in the contig. The cloned gene can be used according to any suitable method known in the art, including, for example, genetic studies, transformation, and the development of novel phenotypes.

Mapped SCN, SDS, or SCN and SDS markers, especially those most closely linked to SCN/SDS resistance can be used to identify homologous clones from soybean genomic libraries, including, for example, soybean genomic libraries made in bacterial artificial chromosomes (BAC), yeast artificial chromosomes (YAC), or P1 bacteriophage. These types of vectors are preferred for positional cloning because they have the capacity to carry larger DNA inserts than possible with other vector technologies. These larger DNA inserts allow the researcher to move physically farther along the chromosome by identifying overlapping clones. Exemplary libraries available for positional cloning efforts in soybean include those described by Meksem et al., 2000; Kanazin et al. (1996) *Proc Natl Acad Sci USA* 93(21):11746-11750; Zhu et al. (1996) *Mol Gen Genet* 252:483-488. Exemplary hybridization methods are disclosed in Examples 4 and 5.

Mapped SCN, SDS, or SCN and SDS markers can be used as DNA probes to hybridize and select homologous genomic clones from such libraries.

Alternatively, the DNA of mapped marker clones are sequenced to design PCR primers that amplify and therefore identify homologous genomic clones from such libraries. Either method is used to identify large-insert soybean clones that is then used to start or finish a contig constructed in chromosome walking to clone an SCN, SDS, or SCN and SDS resistance QTL.

As examples, the positional cloning strategy was successfully used to clone the cystic fibrosis gene in humans (Rommens et al. (1989) *Science* 245:1059-1065), an omega-3 desaturase gene in *Arabidopsis* Arondel et al. (1992) *Science* 258:1353-1355), a protein kinase gene (Pto) conferring fungal resistance in tomato (Martin et al. (1993) *Science* 262:1432-1436), a YAC clone containing the jointless gene that suppresses abscission of flowers and fruit in tomato (Zhang et al. (1994) *Mol Gen Genet* 244:613-621), and sequences comprising the *rhg1* and *Rhg4* genes, disclosed herein.

VI.C.   Mapping Methods

The isolated and purified polynucleotide sequences disclosed herein can also be used in a variety of applications pertaining to mapping SCN and SDS resistance. For example, the isolated polynucleotides disclosed herein are useful in studies of genome organization; in gene structure and organization experiments; in BAC-FISH experiments; in chromosome painting techniques; and in chromosome manipulation.

Thus, in accordance with the present invention, the nucleic acid sequences which encode SCN/SDS resistance polypeptides can also be used to generate hybridization probes which are useful for mapping naturally occurring genomic sequences and/or resistance loci. The sequences can be

mapped to a particular chromosome or to a specific region of the chromosome using well-known techniques. Such techniques include FISH, FACS, or artificial chromosome constructions, such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA

5   libraries as reviewed in Price (1993) *Blood Rev* 7:127-134, and Trask (1991) *Trends Genet* 7:149-154.

FISH (as described in Verma et al. (1988) <u>Human Chromosomes: A Manual of Basic Techniques</u>, Pergamon Press, New York, New York) can be correlated with other physical chromosome mapping techniques and genetic

10   map data. Examples of genetic map data can be found in the 1994 Genome Issue of *Science* (265:1981f). Correlation between the location of the gene encoding SCN, SDS, or both SCN and SDS resistance on a physical chromosomal map and another resistance characteristic, or lack thereof, can help delimit the region of DNA associated with that genetic characteristic. The

15   nucleotide sequences of the subject invention can be used to detect differences in gene sequences between normal, carrier, or susceptible individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis and chromosomal painting using established chromosomal markers can be used for extending genetic maps.

20   Often the placement of a gene on the chromosome of another plant species, such as tomato species or other soybean species, reveals associated markers also found in other plants such as soybeans even if the number or arm of a particular chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides

valuable information to investigators searching for resistance or other genes using positional cloning or other gene discovery techniques. Once the resistance or other gene has been crudely localized by genetic linkage to a particular genomic region, any sequences mapping to that area can represent

5   associated or regulatory genes for further investigation. The nucleotide sequences of the present invention can thus also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier, or susceptible individuals, and to detect gene regulatory sequences (e.g. promoters).

10   Hybridization of the subject DNAs to reference chromosomes can also be performed to give information on relative copy numbers of sequences. Normalization is required to obtain absolute copy number information. One convenient method to do this is to hybridize a probe, for example a cosmid specific to some single locus in the normal haploid genome, to the interphase

15   nuclei of the subject cell or cell population(s) (or those of an equivalent cell or representative cells therefrom, respectively). Quantiation of the hybridization signals in a representative population of such nuclei gives the absolute sequence copy number at that location. Given that information at one locus, the intensity (ratio) information from the hybridization of the subject DNA(s) to

20   the reference condensed chromosomes gives the absolute copy number over the rest of the genome. In practice, use of more than one reference locus can be desirable. In this case, the best fit of the intensity (ratio) data through the reference loci can give a more accurate determination of absolute sequence copy number over the rest of the genome.

Thus, the methods of the present invention can provide information on the absolute copy numbers of substantially all RNA or DNA sequences in subject cell(s) or cell population(s) as a function of the location of those sequences in a reference genome. Additionally, chromosome painting probes

5   can be prepared using the markers and sequence data herein disclosed. Hybridization with one or more of such probes indicates the absolute copy numbers of the sequences to which the probes bind.

Further, when the subject nucleic acid sequences are DNA, the reference copy numbers can be determined by Southern analysis. When the

10   subject nucleic acid sequences are RNA, the reference copy numbers can be determined by Northern analysis.

VI.D.   Assays Kits

In another aspect, the present invention provides assay kits for detecting the presence, in biological samples, of a polynucleotide that encodes a

15   polypeptide of the present invention or of a chromosome bearing a gene or locus of the present invention, the kits comprising a first container that contains a second polynucleotide identical or complementary to a segment of at least 10 contiguous nucleotide bases of, as a preferred example, any of SEQ ID NOs:13 and 16-19.

20   VII.   Recombinant Expression B Expression Cassettes

The term "expression cassette" as used herein means a DNA sequence capable of directing expression of a particular nucleotide sequence in an appropriate host cell, comprising a promoter operably linked to the nucleotide sequence of interest which is operably linked to termination signals. It also

typically comprises sequences required for proper translation of the nucleotide sequence. The expression cassette comprising the nucleotide sequence of interest can be chimeric. The expression cassette can also be one which is naturally occurring but has been obtained in a recombinant form useful for

5   heterologous expression. The expression cassettes can also comprise any further sequences required or selected for the expression of the transgene. Such sequences include, but are not restricted to, transcription terminators, extraneous sequences to enhance expression such as introns, vital sequences, and sequences intended for the targeting of the gene product to specific

10   organelles and cell compartments.

VII.A. Promoters

The expression of the nucleotide sequence in the expression cassette can be under the control of a constitutive promoter or an inducible promoter which initiates transcription only when the host cell is exposed to some

15   particular external stimulus. For bacterial prodcution of a SCN/SDS resistance polypeptide, exemplary promoters include Simian virus 40 early promoter, a long terminal repeat promoter from retrovirus, an actin promoter, a heat shock promoter, and a metallothionein protein. For *in vivo* production of a SCN/SDS resistance polypeptide in plants, exemplary constituitve promoters are derived

20   from the CaMV 35S, rice actin, and maize ubiquitin genes, each described herein below. Exemplary inducible promoters for this purpose include the chemicaly inducible PR-1a promoter and a wound-inducible promoter, also described herein below.

Selected promoters can direct expression in specific cell types (such as leaf epidermal cells, mesophyll cells, root cortex cells) or in specific tissues or organs (roots, leaves or flowers, for example). Exemplary tissue-specific promoters include well-characterized root-, pith-, and leaf-specific promoters, each described herein below.

Depending upon the host cell system utilized, any one of a number of suitable promoters can be used. Promoter selction can be based on expression profile and expression level. The following are non-limiting examples of promoters that can be used in the expression cassettes.

VII.A.1.Constituitive Expression

35S Promoter. The CaMV 35S promoter can be used to drive constituitive gene expression. Construction of the plasmid pCGN1761 is described in the published patent application EP 0 392 225, which is hereby incorporated by reference. pCGN1761 contains the "double" CaMV 35S promoter and the tml transcriptional terminator with a unique *EcoRI* site between the promoter and the terminator and has a pUC-type backbone. A derivative of pCGN1761 is constructed which has a modified polylinker which includes NotI and XhoI sites in addition to the existing *EcoRI* site. This derivative is designated pCGN1761ENX. pCGN1761ENX is useful for the cloning of cDNA sequences or gene sequences (including microbial ORF sequences) within its polylinker for the purpose of their expression under the control of the 35S promoter in transgenic plants. The entire 35S promoter-gene sequence-tml terminator cassette of such a construction can be excised by HindIII, SphI, SalI, and XbaI sites 5' to the promoter and XbaI, BamHI and BglI

sites 3' to the terminator for transfer to transformation vectors such as those described below. Furthermore, the double 35S promoter fragment can be removed by 5' excision with HindIII, SphI, SalI, XbaI, or PstI, and 3' excision with any of the polylinker restriction sites (*EcoRI, NotI* or *XhoI*) for replacement with another promoter.

Actin Promoter. Several isoforms of actin are known to be expressed in most cell types and consequently the actin promoter is a good choice for a constitutive promoter. In particular, the promoter from the rice ActI gene has been cloned and characterized (McElroy et al. (1990) *Plant Cell* 2:163-171). A 1.3 kb fragment of the promoter was found to contain all the regulatory elements required for expression in rice protoplasts. Furthermore, numerous expression vectors based on the ActI promoter have been constructed specifically for use in monocotyledons (McElroy et al. (1991) *Mol Gen Genet* 231:150-160). These incorporate the ActI-intron 1, AdhI 5' flanking sequence and AdhI-intron 1 (from the maize alcohol dehydrogenase gene) and sequence from the CaMV 35S promoter. Vectors showing highest expression were fusions of 35S and ActI intron or the ActI 5' flanking sequence and the ActI intron. Optimization of sequences around the initiating ATG (of the GUS reporter gene) also enhanced expression. The promoter expression cassettes described by McElroy et al. (1991) can be easily modified for gene expression and are particularly suitable for use in monocotyledonous hosts. For example, promoter-containing fragments is removed from the McElroy constructions and used to replace the double 35S promoter in pCGN1761ENX, which is then available for the insertion of specific gene sequences. The fusion genes thus

constructed can then be transferred to appropriate transformation vectors. In a separate report, the rice ActI promoter with its first intron has also been found to direct high expression in cultured barley cells (Chibbar et al. (1993) *Plant Cell Rep* 12:506-509).

5   <u>Ubiquitin Promoter</u>. Ubiquitin is another gene product known to accumulate in many cell types and its promoter has been cloned from several species for use in transgenic plants (e.g. sunflower--Binet et al. (1991) *Plant Science* 79: 87-94 and maize--Christensen et al. (1989) *Plant Molec Biol* 12:619-632). The maize ubiquitin promoter has been developed in transgenic 10   monocot systems and its sequence and vectors constructed for monocot transformation are disclosed in the patent publication EP 0 342 926 which is herein incorporated by reference. Taylor et al. (*1993*) *Plant Cell Rep* 12:491-495 describe a vector (pAHC25) that comprises the maize ubiquitin promoter and first intron and its high activity in cell suspensions of numerous 15   monocotyledons when introduced via microprojectile bombardment. The ubiquitin promoter is suitable for gene expression in transgenic plants, especially monocotyledons. Suitable vectors are derivatives of pAHC25 or any of the transformation vectors described in this application, modified by the introduction of the appropriate ubiquitin promoter and/or intron sequences.

20   <u>VII.A.2.</u>   <u>Inducible Expression</u>

        <u>Chemically Inducible PR-1a Promoter</u>. The double 35S promoter in pCGN1761ENX can be replaced with any other promoter of choice which will result in suitably high expression levels. By way of example, one of the chemically regulatable promoters described in U.S. Patent No. 5,614,395 can

replace the double 35S promoter. The promoter of choice is preferably excised from its source by restriction enzymes, but can alternatively be PCR-amplified using primers that carry appropriate terminal restriction sites. Should PCR-amplification be undertaken, then the promoter should be re-sequenced to

5   check for amplification errors after the cloning of the amplified promoter in the target vector. The chemical/pathogen regulated tobacco PR-1a promoter is cleaved from plasmid pCIB1004 (for construction, see EP 0 332 104, which is hereby incorporated by reference) and transferred to plasmid pCGN1761ENX (Uknes et al. (1992) *The Plant Cell* 4:645-656).

10   pCIB1004 is cleaved with *Nco*I and the resultant 3' overhang of the linearized fragment is rendered blunt by treatment with T4 DNA polymerase. The fragment is then cleaved with *Hind*III and the resultant PR-1a promoter-containing fragment is gel purified and cloned into pCGN1761ENX from which the double 35S promoter has been removed. This is done by

15   cleavage with XhoI and blunting with T4 polymerase, followed by cleavage with HindIII and isolation of the larger vector-terminator containing fragment into which the pCIB1004 promoter fragment is cloned. This generates a pCGN1761ENX derivative with the PR-1a promoter and the tml terminator and an intervening polylinker with unique *Eco*RI and *Not*I sites. The selected coding

20   sequence can be inserted into this vector, and the fusion products (i.e. promoter-gene-terminator) can subsequently be transferred to any selected transformation vector, including those described below. Various chemical regulators can be employed to induce expression of the selected coding sequence in the plants transformed according to the present invention,

including the benzothiadiazole, isonicotinic acid, and salicylic acid compounds

disclosed in U.S. Patent Nos. 5,523,311 and 5,614,395, herein incorporated by

reference.

Wound-Inducible Promoters. Wound-inducible promoters can also be

5   suitable for gene expression. Numerous such promoters have been described

(e.g. Xu et al. (1993) *Plant Molec Biol* 22:573-588; Logemann et al. (1989)

Plant Cell 1:151-158; Rohrmeier & Lehle (1993) *Plant Molec Biol* 22:783-792;

Firek et al. (1993) *Plant Molec Biol* 22:129-142; Warner et al. (1993) *Plant J*

3:191-201) and all are suitable for use with the instant invention. Logemann et

10   al. (1989) describe the 5' upstream sequences of the dicotyledonous potato

wunl gene. Xu et al. (1993) show that a wound-inducible promoter from the

dicotyledon potato (pin2) is active in the monocotyledon rice. Further,

Rohrmeier & Lehle (1993) describe the cloning of the maize Wipl cDNA which

is wound induced and which can be used to isolate the cognate promoter using

15   standard techniques. Similarly, Firek et al. (1993) and Warner et al. (1993)

have described a wound-induced gene from the monocotyledon *Asparagus*

officinalis, which is expressed at local wound and pathogen invasion sites.

Using cloning techniques well known in the art, these promoters can be

transferred to suitable vectors, fused to the genes pertaining to this invention,

20   and used to express these genes at the sites of plant wounding.

VII.A.3.   Tissue-Specific Expression

Root Promoter. Another pattern of gene expression is root expression.

A suitable root promoter is described by de Framond (*1991*) *FEBS* 290:103-106

and also in the published patent application EP 0 452 269, which is herein

incorporated by reference. This promoter is transferred to a suitable vector such as pCGN1761ENX for the insertion of a selected gene and subsequent transfer of the entire promoter-gene-terminator cassette to a transformation vector of interest.

5   Pith Promoter. International Publication No. WO 93/07278, which is herein incorporated by reference, describes the isolation of the maize trpA gene, which is preferentially expressed in pith cells. The gene sequence and promoter extending up to -1726 bp from the start of transcription are presented. Using standard molecular biological techniques, this promoter, or parts thereof,

10   can be transferred to a vector such as pCGN1761 where it can replace the 35S promoter and be used to drive the expression of a foreign gene in a pith-preferred manner. In fact, fragments containing the pith-preferred promoter or parts thereof can be transferred to any vector and modified for utility in transgenic plants.

15   Leaf Promoter. A maize gene encoding phosphoenol carboxylase (PEPC) has been described by Hudspeth & Grula (1989) *Plant Molec Biol* 12:579-589. Using standard molecular biological techniques the promoter for this gene can be used to drive the expression of any gene in a leaf-specific manner in transgenic plants.

20   VII.B. Transcriptional Terminators

A variety of transcriptional terminators are available for use in expression cassettes. These are responsible for the termination of transcription beyond the transgene and its correct polyadenylation. Appropriate transcriptional terminators are those that are known to function in plants and include the CaMV

35S terminator, the tml terminator, the nopaline synthase terminator and the pea rbcS E9 terminator. These can be used in both monocotyledons and dicotyledons.

VII.C. Sequences for the Enhancement or Regulation of Expression

5   Numerous sequences have been found to enhance gene expression from within the transcriptional unit and these sequences can be used in conjunction with the genes of this invention to increase their expression in transgenic plants.

If desired, modifications around the cloning sites can be made by the

10   introduction of sequences that can enhance translation. This is particularly useful when overexpression is desired. For example, pCGN1761ENX can be modified by optimization of the translational initiation site as disclosed in U.S. Patent No. 5,639,949, incorporated herein by reference.

Various intron sequences have been shown to enhance expression,

15   particularly in monocotyledonous cells. For example, the introns of the maize AdhI gene have been found to significantly enhance the expression of the wild-type gene under its cognate promoter when introduced into maize cells. Intron 1 was found to be particularly effective and enhanced expression in fusion constructs with the chloramphenicol acetyltransferase gene (Callis et al.

20   (1987) Genes Develop 1:1183-1200). In the same experimental system, the intron from the maize bronze1 gene had a similar effect in enhancing expression. Intron sequences have been routinely incorporated into plant transformation vectors, typically within the non-translated leader.

A number of non-translated leader sequences derived from viruses are also known to enhance expression, and these are particularly effective in dicotyledonous cells. Specifically, leader sequences from Tobacco Mosaic Virus (TMV, the "W-sequence"), Maize Chlorotic Mottle Virus (MCMV), and

5 Alfalfa Mosaic Virus (AMV) have been shown to be effective in enhancing expression (e.g. Gallie et al. *(1987) Nucl Acids Res* 15:8693-8711; Skuzeski et al. (1990) *Plant Molec Biol* 15:65-79).

VII.D. Targeting of the Gene Product Within the Cell

Various mechanisms for targeting gene products are known to exist in

10 plants and the sequences controlling the functioning of these mechanisms have been characterized in some detail. For example, the targeting of gene products to the chloroplast is controlled by a signal sequence found at the amino terminal end of various proteins which is cleaved during chloroplast import to yield the mature protein (e.g. Comai et al. (1988) *J Biol Chem* 263:15104-15109). These

15 signal sequences can be fused to heterologous gene products to effect the import of heterologous products into the chloroplast (van den Broeck et al. (1985) *Nature* 313:358-363). DNA encoding for appropriate signal sequences can be isolated from the 5' end of the cDNAs encoding the RUBISCO protein, the CAB protein, the EPSP synthase enzyme, the GS2 protein and many other

20 proteins which are known to be chloroplast localized. See also, U.S. Patent No. 5,639,949, herein incorporated by reference.

Other gene products are localized to other organelles such as the mitochondrion and the peroxisome (e.g. Unger et al. (1989) *Plant Molec Biol* 13:411-418). The cDNAs encoding these products can also be manipulated to

effect the targeting of heterologous gene products to these organelles. Examples of such sequences are the nuclear-encoded ATPases and specific aspartate amino transferase isoforms for mitochondria. Targeting cellular protein bodies has been described by Rogers et al. (1989) *Proc Natl Acad Sci*

5   *USA* 82:6512-6516).

In addition, sequences have been characterized which cause the targeting of gene products to other cell compartments. Amino terminal sequences are responsible for targeting to the ER, the apoplast, and extracellular secretion from aleurone cells (Koehler & Ho (1990) *Plant Cell*

10   2:769-783). Additionally, amino terminal sequences in conjunction with carboxy terminal sequences are responsible for vacuolar targeting of gene products (Shinshi et al. (1990) *Plant Molec Biol* 14:357-368).

By the fusion of the appropriate targeting sequences described above to transgene sequences of interest, it is possible to direct the transgene product to

15   any organelle or cell compartment. For chloroplast targeting, for example, the chloroplast signal sequence from the RUBISCO gene, the CAB gene, the EPSP synthase gene, or the GS2 gene is fused in frame to the amino terminal ATG of the transgene. The signal sequence selected should include the known cleavage site, and the fusion constructed should take into account any amino

20   acids after the cleavage site which are required for cleavage. In some cases this requirement can be fulfilled by the addition of a small number of amino acids between the cleavage site and the transgene ATG or, alternatively, replacement of some amino acids within the transgene sequence. Fusions constructed for chloroplast import can be tested for efficacy of chloroplast

uptake by *in vitro* translation of *in vitro* transcribed constructions followed by *in vitro* chloroplast uptake using techniques described by Bartlett et al. (1982) in Methods in Chloroplast Molecular Biology, Edelmann et al. (Eds.), pp 1081-1091, Elsevier and Wasmann et al. (1986) *Mol Gen Genet* 205:446-453.

5   These construction techniques are well known in the art and are equally applicable to mitochondria and peroxisomes.

The above-described mechanisms for cellular targeting can be utilized not only in conjunction with their cognate promoters, but also in conjunction with heterologous promoters so as to effect a specific cell-targeting goal under the

10   transcriptional regulation of a promoter that has an expression pattern different to that of the promoter from which the targeting signal derives.

VIII.   Recombinant Expression B Vectors

Suitable expression vectors which can be used include, but are not limited to, the following vectors or their derivatives: human or animal viruses

15   such as vaccinia virus or adenovirus, yeast vectors, bacteriophage vectors (e.g., lambda phage), and plasmid and cosmid DNA vectors.

Numerous vectors available for plant transformation are known to those of ordinary skill in the plant transformation arts, and the genes pertinent to this invention can be used with any such vectors. Exemplary vectors include

20   pCIB200, pCIB2001, pCIB10, pCIB3064, pSOG19, and pSOG35, each described herein below. The selection of vector will depend upon the preferred transformation technique and the target species for transformation.

VIII.A. _Agrobacterium_ Transformation Vectors.

Many vectors are available for transformation using _Agrobacterium tumefaciens._ These typically carry at least one T-DNA border sequence and include vectors such as pBIN19 (Bevan (1984) _Nucl Acids Res_ 12:8711-8721) and pXYZ. Below, the construction of two typical vectors suitable for _Agrobacterium_ transformation is described.

pCIB200 and pCIB2001. The binary vectors pcIB200 and pCIB2001 are used for the construction of recombinant vectors for use with _Agrobacterium_ and are constructed in the following manner. pTJS75kan is created by NarI digestion of pTJS75 (Schmidhauser & Helinski (1985) _J Bacteriol_ 164:446-455) allowing excision of the tetracycline-resistance gene, followed by insertion of an AccI fragment from pUC4K carrying an NPTII (Messing & Vierra (1982) _Gene_ 19:259-268; Bevan et al. (1983) _Nature_ 304:184-187; McBride et al. (1990) _Plant Molecular Biology_ 14:266-276). XhoI linkers are ligated to the _EcoRV_ fragment of PCIB7 which contains the left and right T-DNA borders, a plant selectable nos/nptII chimeric gene and the pUC polylinker (Rothstein et al. (1987) _Gene_ 53:153-161), and the XhoI-digested fragment are cloned into SalI-digested pTJS75kan to create pCIB200 (see also EP 0 332 104, herein incorporated by reference).

pCIB200 contains the following unique polylinker restriction sites: _EcoRI_, _SstI_, _KpnI_, _BglII_, _XbaI_, and _SalI_. pCIB2001 is a derivative of pCIB200 created by the insertion into the polylinker of additional restriction sites. Unique restriction sites in the polylinker of pCIB2001 are _EcoRI_, _SstI_, _KpnI_, _BglII_, _XbaI_, _SalI_, _MluI_, _BclI_, _AvrII_, _ApaI_, _HpaI_, and _StuI_. pCIB2001, in addition to containing

these unique restriction sites also has plant and bacterial kanamycin selection,
left and right T-DNA borders for *Agrobacterium*-mediated transformation, the
RK2-derived trfA function for mobilization between *E. coli* and other hosts, and
the OriT and OriV functions also from RK2. The pCIB2001 polylinker is suitable
5   for the cloning of plant expression cassettes containing their own regulatory
signals.

pCIB10 and Hygromycin Selection Derivatives thereof. The binary
vector pCIB10 contains a gene encoding kanamycin resistance for selection in
plants and T-DNA right and left border sequences and incorporates sequences
10   from the wide host-range plasmid pRK252 allowing it to replicate in both *E. coli*
and *Agrobacterium*.   Its construction is described by Rothstein et al. (1987).
Various derivatives of pCIB10 are constructed which incorporate the gene for
hygromycin B phosphotransferase described by Gritz et al. (1983) *Gene*
25:179-188. These derivatives enable selection of transgenic plant cells on
15   hygromycin only (pCIB743), or hygromycin and kanamycin (pCIB715,
pCIB717).

VIII.B. Other Plant Transformation Vectors

Transformation without the use of *Agrobacterium tumefaciens*
circumvents the requirement for T-DNA sequences in the chosen
20   transformation vector and consequently vectors lacking these sequences can
be utilized in addition to vectors such as the ones described above which
contain T-DNA sequences. Transformation techniques that do not rely on
Agrobacterium include transformation via particle bombardment, protoplast
uptake (e.g. PEG and electroporation) and microinjection. The choice of vector

depends largely on the preferred selection for the species being transformed. Below, the construction of typical vectors suitable for non-*Agrobacterium* transformation is described.

pCIB3064. pCIB3064 is a pUC-derived vector suitable for direct gene

5 transfer techniques in combination with selection by the herbicide basta (or phosphinothricin). The plasmid pCIB246 comprises the CaMV 35S promoter in operational fusion to the *E. coli* GUS gene and the CaMV 35S transcriptional terminator and is described in the Internation Publication No. WO 93/07278. The 35S promoter of this vector contains two ATG sequences 5' of the start

10 site. These sites are mutated using standard PCR techniques in such a way as to remove the ATGs and generate the restriction sites SspI and PvuII. The new restriction sites are 96 and 37 bp away from the unique SalI site and 101 and 42 bp away from the actual start site. The resultant derivative of pCIB246 is designated pCIB3025.

15 The GUS gene is then excised from pCIB3025 by digestion with SalI and SacI, the termini rendered blunt and religated to generate plasmid pCIB3060. The plasmid pJIT82 is obtained from the John Innes Centre, Norwich and the a 400 bp SmaI fragment containing the bar gene from *Streptomyces viridochromogenes* is excised and inserted into the HpaI site of pCIB3060

20 (Thompson et al. (1987) *EMBO J* 6:2519-2523). This generated pCIB3064, which comprises the bar gene under the control of the CaMV 35S promoter and terminator for herbicide selection, a gene for ampicillin resistance (for selection in *E. coli*) and a polylinker with the unique sites SphI, PstI, HindIII, and BamHI.

This vector is suitable for the cloning of plant expression cassettes containing their own regulatory signals.

pSOG19 and pSOG35. pSOG35 is a transformation vector that utilizes the *E. coli* gene dihydrofolate reductase (DFR) as a selectable marker conferring resistance to methotrexate. PCR is used to amplify the 35S promoter (-800 bp), intron 6 from the maize Adh1 gene (-550 bp) and 18 bp of the GUS untranslated leader sequence from pSOG10. A 250-bp fragment encoding the *E. coli* dihydrofolate reductase type II gene is also amplified by PCR and these two PCR fragments are assembled with a SacI-PstI fragment from pB1221 (Clontech, Palo Alto, California) which comprises the pUC19 vector backbone and the nopaline synthase terminator. Assembly of these fragments generates pSOG19 which contains the 35S promoter in fusion with the intron 6 sequence, the GUS leader, the DHFR gene and the nopaline synthase terminator. Replacement of the GUS leader in pSOG19 with the leader sequence from Maize Chlorotic Mottle Virus (MCMV) generates the vector pSOG35. pSOG19 and pSOG35 carry the pUC gene for ampicillin resistance and have HindIII, SphI, PstI and *Eco*RI sites available for the cloning of foreign substances.

VIII.C. Selectable Markers

For certain target species, different antibiotic or herbicide selection markers can be preferred. Selection markers used routinely in transformation include the nptII gene, which confers resistance to kanamycin and related antibiotics (Messing & Vierra (1982) *Gene* 19:259-268; Bevan et al., 1983), the bar gene, which confers resistance to the herbicide phosphinothricin (White et

al. (1990) *Nucl Acids Res* 18:1062; Spencer et al. (1990) *Theor Appl Genet* 79:625-631), the hph gene, which confers resistance to the antibiotic hygromycin (Blochlinger & Diggelmann (1984) *Mol Cell Biol* 4:2929-2931), the dhfr gene, which confers resistance to methatrexate (Bourouis et al., (1983)

5   *EMBO J* 2(7):1099-1104), and the EPSPS gene, which confers resistance to glyphosate (U.S. Patent Nos. 4,940,935 and 5,188,642).

IX.   Recombinant Expression B Host Cells

The term "host cell", as used herein, refers to a cell into which a heterologous nucleic acid molecule has been introduced. Transformed cells,

10   tissues, or organisms are understood to encompass not only the end product of a transformation process, but also transgenic progeny thereof. A host cell strain can be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. For example, different host cells have characteristic and specific

15   mechanisms for the translational and post-translational processing and modification (e.g., glycosylation, phosphorylation of proteins). Appropriate cell lines or host systems can be chosen to ensure the desired modification and processing of the foreign protein expressed. Expression in a bacterial system can be used to produce a non-glycosylated core protein product. Expression in

20   yeast will produce a glycosylated product. Expression in plant cells can be used to ensure "native" glycosylation of a heterologous protein.

The present invention provides methods for recombinant expression of SCN/SDS resistance genes in plants by the construction of transgenic plants. The phrase "a plant, or parts thereof" as used herein shall mean an entire plant;

and shall mean the individual parts thereof, including but not limited to seeds, leaves, stems, and roots, as well as plant tissue cultures. Transgenic plants of the present invention are understood to encompass not only the end product of a transformation method, but also transgenic progeny thereof. The term

5   "converted plant" as used herein shall mean any plant (1) having resistance to SDS or resistance to SCN and (2) and was derived by genetic selection employing RFLP, RADP, AFLP, or microsatellite (SSR) data for at least one of the loci herein defined.

Preferably, the plant is a soybean plant. However, disease resistance

10   can be conferred to a wide variety of plant cells, including those of gymnosperms, monocots, and dicots. Although the gene can be inserted into any plant cell falling within these broad classes, it is particularly useful in crop plant cells, such as rice, wheat, barley, rye, corn, potato, carrot, sweet potato, sugar beet, bean, pea, chicory, lettuce, cabbage, cauliflower, broccoli, turnip,

15   radish, spinach, asparagus, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, quince, melon, plum, cherry, peach, nectarine, apricot, strawberry, grape, raspberry, blackberry, pineapple, avocado, papaya, mango, banana, tobacco, tomato, sorghum and sugarcane.

20   X.   Recombinant Expression B Transfection and Transformation Methods

Expression constructs are transfected into a host cell by a standard method suitable for the selected host, including electroporation, calcium phosphate precipitation, DEAE-Dextran transfection, liposome-mediated transfection, infection using a retrovirus, transposon-mediated transfer, and

particle bombardment techniques. The SCN/SDS resistance gene-encoding nucleotide sequence carried in the expression construct can be stably integrated into the genome of the host or it can be present as an extrachromosomal molecule. Below are descriptions of representative techniques for transforming both dicotyledonous and monocotyledonous plants.

X.A. Transformation of Dicotyledons

Transformation techniques for dicotyledons are well known in the art and include *Agrobacterium*-based techniques and techniques that do not require *Agrobacterium*. Non-*Agrobacterium* techniques involve the uptake of exogenous genetic material directly by protoplasts or cells. This can be accomplished by PEG or electroporation mediated uptake, particle bombardment-mediated delivery, or microinjection. Examples of these techniques are described by Paszkowski et al. (1984) *EMBO J* 3:2717-2722; Potrykus et al. (1985) *Mol Gen Genet* 199:169-177; Reich et al. (1986) *Biotechnology* 4:1001-1004; and Klein et al. (1987) *Nature* 327:70-73. In each case the transformed cells are regenerated to whole plants using standard techniques known in the art.

Agrobacterium-mediated transformation is a preferred technique for transformation of dicotyledons because of its high efficiency of transformation and its broad utility with many different species. *Agrobacterium* transformation typically involves the transfer of the binary vector carrying the foreign DNA of interest (e.g. pCIB200 or pCIB2001) to an appropriate *Agrobacterium* strain, which can depend of the complement of vir genes carried by the host *Agrobacterium* strain either on a co-resident Ti plasmid or chromosomally (e.g.

strain CIB542 for pCIB200 and pCIB2001 (Uknes et al. (1993) *Plant Cell* 5:159-169). The transfer of the recombinant binary vector to *Agrobacterium* is accomplished by a triparental mating procedure using *E. coli* carrying the recombinant binary vector, a helper *E. coli* strain which carries a plasmid such

5   as pRK2013 and which is able to mobilize the recombinant binary vector to the target *Agrobacterium* strain. Alternatively, the recombinant binary vector can be transferred to *Agrobacterium* by DNA transformation (Hofgen & Willmitzer (1988) *Nucl Acids Res* 16:9877).

Transformation of the target plant species by recombinant

10   *Agrobacterium* usually involves co-cultivation of the *Agrobacterium* with explants from the plant and follows protocols well known in the art. Transformed tissue is regenerated on selectable medium carrying the antibiotic or herbicide resistance marker present between the binary plasmid T-DNA borders.

Another approach to transforming plant cells with a gene involves

15   propelling inert or biologically active particles at plant tissues and cells. This technique is disclosed in U.S. Patent Nos. 4,945,050, 5,036,006, and 5,100,792. Generally, this procedure involves propelling inert or biologically active particles at the cells under conditions effective to penetrate the outer surface of the cell and afford incorporation within the interior thereof. When

20   inert particles are utilized, the vector can be introduced into the cell by coating the particles with the vector containing the desired gene. Alternatively, the target cell can be surrounded by the vector so that the vector is carried into the cell by the wake of the particle. Biologically active particles (e.g., dried yeast

cells, dried bacterium or a bacteriophage, each containing DNA sought to be introduced) can also be propelled into plant cell tissue.

X.B. Transformation of Monocotyledons

Transformation of most monocotyledon species has now also become
5   routine. Preferred techniques include direct gene transfer into protoplasts using PEG or electroporation techniques, and particle bombardment into callus tissue. Transformations can be undertaken with a single DNA species or multiple DNA species (i.e. co-transformation) and both these techniques are suitable for use with this invention. Co-transformation can have the advantage
10   of avoiding complete vector construction and of generating transgenic plants with unlinked loci for the gene of interest and the selectable marker, enabling the removal of the selectable marker in subsequent generations, should this be regarded desirable. However, a disadvantage of the use of co-transformation is the less than 100% frequency with which separate DNA species are integrated
15   into the genome (Schocher et al. (1986) *Biotechnology* 4:1093-1096).

Patent Application Nos. EP 0 292 435, EP 0 392 225, and International Publication No. WO 93/07278 describe techniques for the preparation of callus and protoplasts from an elite inbred line of maize, transformation of protoplasts using PEG or electroporation, and the regeneration of maize plants from
20   transformed protoplasts. Gordon-Kamm et al. (1990) *Plant Cell* 2:603-618 and Fromm et al. (1990) *Biotechnology* 8:833-839 have published techniques for transformation of A188-derived maize line using particle bombardment. Furthermore, International Publication No. WO 93/07278 and Koziel et al. (1993) *Biotechnology* 11:194-200 describe techniques for the transformation of

elite inbred lines of maize by particle bombardment. This technique utilizes immature maize embryos of 1.5-2.5 mm length excised from a maize ear 14-15 days after pollination and a PDS-1000He BIOLISTICS® device for bombardment.

5   Transformation of rice can also be undertaken by direct gene transfer techniques utilizing protoplasts or particle bombardment. Protoplast-mediated transformation has been described for Japonica-types and Indica-types (Zhang et al. (1988) *Plant Cell Rep* 7:379-384; Shimamoto et al. (1989) *Nature* 338:274-277; Datta et al. (1990) *Biotechnology* 8:736-740). Both types are also

10   routinely transformable using particle bombardment (Christou et al. (1991) *Biotechnology* 9:957-962). Furthermore, Internation Publication Number WO 93/21335 describes techniques for the transformation of rice via electroporation. Patent Application EP 0 332 581 describes techniques for the generation, transformation and regeneration of *Pooideae* protoplasts. These techniques

15   allow the transformation of *Dactylis* and wheat. Furthermore, wheat transformation has been described by Vasil et al. (1992) *Biotechnology* 10:667-674 using particle bombardment into cells of type C long-term regenerable callus, and also by Vasil et al. (1993) *Biotechnology* 11:1553-1558 and Weeks et al. (1993) *Plant Physiol* 102:1077-1084 using particle

20   bombardment of immature embryos and immature embryo-derived callus. A preferred technique for wheat transformation, however, involves the transformation of wheat by particle bombardment of immature embryos and includes either a high sucrose or a high maltose step prior to gene delivery. Prior to bombardment, any number of embryos (0.75-1 mm in length) are plated

onto MS medium with 3% sucrose (Murashiga & Skoog (1962) *Physiologia Plantarum* 15:473-497) and 3 mg/l 2,4-D for induction of somatic embryos, which is allowed to proceed in the dark. On the chosen day for bombardment, embryos are removed from the induction medium and placed onto the

5   osmoticum (i.e. induction medium with sucrose or maltose added at the desired concentration, typically 15%). The embryos are allowed to plasmolyze for 2-3 h and are then bombarded. Twenty embryos per target plate is typical, although not critical.

An appropriate gene-carrying plasmid (such as pCIB3064 or pSG35) is

10   precipitated onto micrometer size gold particles using standard procedures. Each plate of embryos is shot with the DuPont BIOLISTICS® helium device using a burst pressure of about 1000 psi using a standard 80 mesh screen. After bombardment, the embryos are placed back into the dark to recover for about 24 hours (still on osmoticum). After 24 hours, the embryos are removed

15   from the osmoticum and placed back onto induction medium where they stay for about a month before regeneration. Approximately one month later the embryo explants with developing embryogenic callus are transferred to regeneration medium (MS+1 mg/liter NAA, 5 mg/liter GA), further containing the appropriate selection agent (10 mg/l basta in the case of pCIB3064 and 2 mg/l methotrexate

20   in the case of pSOG35). After approximately one month, developed shoots are transferred to larger sterile containers known as "GA7s" which contain half-strength MS, 2% sucrose, and the same concentration of selection agent.

More recently, tranformation of monocotyledons using *Agrobacterium* has been described. See WO 94/00977 and U.S. Patent No. 5,591,616, both of which are incorporated herein by reference.

XI. Antibodies

5   The present invention also provides an antibody immunoreactive with an SCN/SDS resistance polypeptide. The term "antibody" indicates an immunoglobulin protein, or functional portion thereof, including a polyclonal antibody, a monoclonal antibody, a chimeric antibody, a single chain antibody, Fab fragments, and an Fab expression library. "Functional portion" refers to the

10   part of the protein that binds a molecule of interest. In a preferred embodiment, an antibody of the invention is a monoclonal antibody. Techniques for preparing and characterizing antibodies are well known in the art (See, *e.g.*, Harlow and Lane (1988). A monoclonal antibody of the present invention can be readily prepared through use of well-known techniques such as the hybridoma

15   techniques exemplified in U.S. Patent No 4,196,265 and the phage-displayed techniques disclosed in U.S. Patent No. 5,260,203.

The phrase "specifically (or selectively) binds to an antibody", or "specifically (or selectively) immunoreactive with", when referring to a protein or peptide, refers to a binding reaction which is determinative of the presence of

20   the protein in a heterogeneous population of proteins and other biological materials. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein and do not show significant binding to other proteins present in the sample. Specific binding to an antibody under such conditions can require an antibody that is selected for its specificity for a

particular protein. For example, antibodies raised to a protein with an amino acid sequence encoded by the nucleic acid sequence of SEQ ID No:13 can be selected to obtain antibodies specifically immunoreactive with that protein and not with unrelated proteins.

5   The use of a molecular cloning approach to generate antibodies, particularly monoclonal antibodies, and more particularly single chain monoclonal antibodies, are also provided. The production of single chain antibodies has been described in the art. See, e.g., U.S. Patent No. 5,260,203. For this approach, combinatorial immunoglobulin phagemid libraries are

10   prepared from RNA isolated from the spleen of the immunized animal, and phagemids expressing appropriate antibodies are selected by panning on endothelial tissue. The advantages of this approach over conventional hybridoma techniques are that approximately 10^4 times as many antibodies can be produced and screened in a single round, and that new specificities are

15   generated by heavy (H) and light (L) chain combinations in a single chain, which further increases the chance of finding appropriate antibodies. Thus, an antibody of the present invention, or a "derivative" of an antibody of the present invention, pertains to a single polypeptide chain binding molecule which has binding specificity and affinity substantially similar to the binding specificity and

20   affinity of the light and heavy chain aggregate variable region of an antibody described herein.

The term "immunochemical reaction", as used herein, refers to any of a variety of immunoassay formats used to detect antibodies specifically bound to a particular protein, including but not limited to, competitive and non-competitive

assay systems using techniques such as radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitin reactions, immunodiffusion assays, *in situ* immunoassays (e.g., using colloidal gold, enzyme or radioisotope labels),

5   western blots, precipitation reactions, agglutination assays (*e.g.*, gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. <u>See</u> Harlow and Lane (1988) for a description of immunoassay formats and conditions.

10   <u>XII.</u>   <u>Method for Detecting a SCN/SDS Resistance Polypeptide</u>

In another aspect of the invention, a method is provided for detecting a level of SCN/SDS resistance polypeptide using an antibody that specifically recognizes a SCN/SDS resistance polypeptide, or portion thereof. In a preferred embodiment, biological samples from an experimental plant and a

15   control plant are obtained, and SCN/SDS resistance polypeptide is detected in each sample by immunochemical reaction with the SCN/SDS resistance polypeptide antibody. More preferably, the antibody recognizes amino acids of SEQ ID NO:14 and is prepared according to a method of the present invention for producing such an antibody.

20   In one embodiment, a SCN/SDS resistance polypeptide antibody is used to screen a biological sample for the presence of a SCN/SDS resistance polypeptide. A biological sample to be screened can be a biological fluid such as extracellular or intracellular fluid, or a cell or tissue extract or homogenate. A biological sample can also be an isolated cell (*e.g.*, in culture) or a collection of

cells such as in a tissue sample. A tissue sample can be suspended in a liquid medium or fixed onto a solid support such as a microscope slide. In accordance with a screening assay method, a biological sample is exposed to an antibody immunoreactive with an SCN/SDS resistance polypeptide whose presence is

5   being assayed, and the formation of antibody-polypeptide complexes is detected. Techniques for detecting such antibody-antigen conjugates or complexes are well known in the art and include but are not limited to centrifugation, affinity chromatography and the like, and binding of a labeled secondary antibody to the antibody-candidate receptor complex.

10   XIII.   Identification of Modulators of SCN/SDS Resistance

The present invention further discloses a method for identifying a compound that modulates SCN/SDS resistance. As used herein, the terms "candidate substance" and "candidate compound" are used interchangeably and refer to a substance that is believed to interact with another moiety, wherein a

15   biological activity is modulated. For example, a representative candidate compound is believed to interact with a complete, or a fragment of, a SCN/SDS resistance polypeptide, and which can be subsequently evaluated for such an interaction. Exemplary candidate compounds that can be investigated using the methods of the present invention include, but are not restricted to, compounds

20   that confer SCN/SDS resistance, viral epitopes, peptides, enzymes, enzyme substrates, co-factors, lectins, sugars, oligonucleotides or nucleic acids, oligosaccharides, proteins, chemical compounds small molecules, and monoclonal antibodies. A candidate compound to be tested by these methods

can be a purified molecule, a homogenous sample, or a mixture of molecules or compounds.

As used herein, the term "modulate" means an increase, decrease, or other alteration of any or all chemical and biological activities or properties of a wild-type SCN/SDS resistance polypeptide, preferably a SCN/SDS resistance polypeptide of SEQ ID NO:14. Preferably, a SCN/SDS resistance modulator is an agonist of SCN/SDS resistance protein activity. As used herein, the term "agonist" means a substance that supplements or potentiates the biological activity of a functional SCN/SDS resistance protein.

In accordance with the present invention there is also provided a rapid and high throughput screening method that relies on the methods described above. This screening method comprises separately contacting each compound with a plurality of substantially identical samples. In such a screening method the plurality of samples preferably comprises more than about 10^4 samples, or more preferably comprises more than about 5×10^4 samples. In an alternative high-throughput strategy, each sample can be contacted with a plurality of candidate compounds.

XIII.A. Methods for Identifying Modulators of SCN/SDS Resistance Gene Expression

The nucleic acid sequences of the present invention can be used to identify regulators of SCN/SDS resistance polypeptide gene expression. Several molecular cloning strategies can be used to identify substances that specifically bind SCN/SDS resistance polypeptide cis-regulatory elements. A preferred promoter region to be used in such assays is an SCN/SDS resistance

polypeptide promoter region from soybean, more preferably the promoter region includes some or all amino acids of SEQ ID NO:14.

In one embodiment, a cDNA library in an expression vector, such as the lambda-gt11 vector, can be screened for cDNA clones that encode an
5   SCN/SDS resistance polypeptide regulatory element DNA-binding activity by probing the library with a labeled SCN/SDS resistance polypeptide DNA fragment, or synthetic oligonucleotide (Singh et al. (1989) *Biotechniques* 7:252-261). Preferably the nucleotide sequence selected as a probe has already been demonstrated as a protein binding site using a protein-DNA binding assay
10   described above.

In another embodiment, transcriptional regulatory proteins are identified using the yeast one-hybrid system (Luo et al. (1996) *Biotechniques* 20(4):564-568; Vidal et al. (1996) *Proc Natl Acad Sci USA* 93(19):10315-10320; Li and Herskowitz (1993) *Science* 262:1870-1874). In this case, a cis-regulatory
15   element of a SCN/SDS resistance gene is operably fused as an upstream activating sequence (UAS) to one, or typically more, yeast reporter genes such as the lacZ gene, the URA3 gene, the LEU2 gene, the HIS3 gene, or the LYS2 gene, and the reporter gene fusion construct(s) is inserted into an appropriate yeast host strain. It is expected that the reporter genes are not transcriptionally
20   active in the engineered yeast host strain, for lack of a transcriptional activator protein to bind the UAS derived from the SCN/SDS resistance gene promoter region. The engineered yeast host strain is transformed with a library of cDNAs inserted in a yeast activation domain fusion protein expression vector, *e.g.* pGAD, where the coding regions of the cDNA inserts are fused to a functional

yeast activation domain coding segment, such as those derived from the GAL4 or VP16 activators. Transformed yeast cells that acquire a cDNA encoding a protein that binds a cis-regulatory element of a SCN/SDS resistance gene can be identified based on the concerted activation the reporter genes, either by

5 genetic selection for prototrophy (e.g. LEU2, HIS3, or LYS2 reporters) or by screening with chromogenic substrates (lacZ reporter) by methods known in the art.

The present invention also provides an *in vivo* assay for discovery of modulators of SCN/SDS resistance gene expression. In this case, a transgenic

10 plant is made such that a transgene comprising a SCN/SDS resistance gene promoter and a reporter gene is expressed and a level of reporter gene expression is assayable. Such transgenic animals can be used for the identification of compounds that are effective in modulating SCN/SDS resistance gene expression.

15 *In vitro* or *in vivo* screening approaches may survey more than one modulatable transcriptional regulatory sequence simultaneously.

XIII.B. Methods for Identifying Modulators of SCN/SDS Resistance Polypeptides

According to the method, a SCN/SDS resistance polypeptide is exposed

20 to a plurality of candidate substances, and binding of a candidate substance to the SCN/SDS resistance polypeptide is assayed. A compound is selected that demonstrates specific binding to the SCN/SDS resistance polypeptide. Preferably, the SCN/SDS resistance polypeptide used in the binding assay of the method includes some or all amino acids of SEQ ID NO:14.

The term "binding" refers to an affinity between two molecules, for example, a ligand and a receptor, means a preferential binding of one molecule for another in a mixture of molecules. The binding of the molecules can be considered specific if the binding affinity is about 1×10^4 M^{-1} to about 1×10^6 M^{-1}

5   or greater. Binding of two molecules also encompasses a quality or state of mutual action such that an activity of one protein or compound on another protein is inhibitory (in the case of an antagonist) or enhancing (in the case of an agonist).

Several techniques can be used to detect interactions between a protein

10   and a chemical ligand without employing an *in vivo* ligand. Representative methods include, but are not limited to, fluorescence correlation spectroscopy, surface-enhanced laser desorption/ionization, and biacore technology, each described herein below. These methods are amenable to automated, high-throughput screening.

15   <u>Fluorescence Correlation Spectroscopy (FCS)</u>. FCS measures the average diffusion rate of a fluorescent molecule within a small sample volume (Madge et al. (1972) *Phys Re Lett* 29:705-708, Maiti et al. (1997) *Proc Natl Acad Sci USA*, 94:11753-11757). The sample size can be as low as 10^3 fluorescent molecules and the sample volume as low as the cytoplasm of a

20   single bacterium. The diffusion rate is a function of the mass of the molecule and decreases as the mass increases. FCS can therefore be applied to protein-ligand interaction analysis by measuring the change in mass and therefore in diffusion rate of a molecule upon binding. In a typical experiment, the target to be analyzed is expressed as a recombinant protein with a sequence tag, such

as a poly-histidine sequence, inserted at the N-terminus or C-terminus. The target protein is expressed in *E. coli*, yeast, or plant cells. The protein is purified by chromatography. For example, the poly-histidine tag can be used to bind the expressed protein to a metal chelate column such as Ni^{2+} chelated on

5   iminodiacetic acid agarose. The protein is then labeled with a fluorescent tag such as carboxytetramethylrhodamine or BODIPY^{TM} (Molecular Probes, Eugene, Oregon). The protein is then exposed in solution to a candidate compound, and its diffusion rate is determined by FCS , using for example, instrumentation available from Carl Zeiss, Inc. (Thornwood, New York). Ligand

10   binding is determined by changes in the diffusion rate of the protein.

Surface-Enhanced Laser Desorption/Ionization (SELDI). SELDI can be used in combination with a time-of-flight mass spectrometer (TOF) to provide a means to rapidly analyze molecules retained on a chip (Hutchens and Yip (1993) *Rapid Commun Mass Spectrom* 7:576-580). It can be applied to ligand-

15   protein interaction analysis by covalently binding the target protein on the chip and using mass spectroscopy to analyze the small molecules that bind to the target protein (Worrall et al. (1998) *Anal Biochem* 70:750-756). In a typical experiment, the target to be analyzed is recombinantly expressed, optionally with a tag, such as poly-histidine, to facilitate purification and handling. The

20   purified protein is bound to the SELDI chip either by utilizing the poly-histidine tag or by other interaction such as ion exchange or hydrophobic interaction. The chip thus prepared is then exposed to a candidate compound via, for example, a delivery system able to pipet the ligands in a sequential manner (autosampler). The chip is then washed in buffers of increasing stringency, for

example a series of buffer solutions containing incrementally increasing ionic strength. After each wash, the bound material is analyzed by SELDI-TOF. Compounds that specifically bind the target are identified by elution in high stringency wahes.

5   Biacore. Biacore technology utilizes changes in the refractive index at the surface layer upon binding of a ligand to a protein immobilized on the layer. In this system, a collection of small ligands is injected sequentially in a 2-5 microliter cell, wherein the protein is immobilized within the cell. Binding is detected by surface plasmon resonance (SPR) of laser light refracting from the

10   surface. In general, the refractive index change for a given change of mass concentration at the surface layer is practically the same for all proteins and peptides, allowing a single method to be applicable for any protein (Liedberg et al. (1983) *Sensors Actuators* 4:299-304; Malmquist (1993) *Nature* 361:186-187). In a typical experiment, the target protein to be analyzed is recombinantly

15   expressed an purified according to standard methods. It is bound to the Biacore chip either by utilizing a poly-histidine tag or by other interaction such as ion exchange or hydrophobic interaction. The chip thus prepared is then exposed to a candidate compound via the delivery system incorporated in the instruments sold by Biacore (Uppsala, Sweden) to pipet the ligands in a

20   sequential manner (autosampler). The SPR signal on the chip is recorded and changes in the refractive index indicate an interaction between the immobilized target and the ligand. Analysis of the signal kinetics on rate and off rate allows the discrimination between non-specific and specific interaction.

Rational Drug Design. Similarly, the knowledge of the structure a native SCN/SDS resistance polypeptide provides an approach for rational drug design. The structure of an SCN/SDS resistance polypeptide can be determined by X-ray crystallography or by computational algorithms that generate three-

5   dimensional representations. See Huang et al. (2000) and Saqi et al. (1999) Computer models can further predict binding of a protein structure to various substrate molecules, that can be synthesized and tested. Additional drug design techniques are described in U.S. Patent Nos. 5,834,228 and 5,872,011.

XIV.   Modulation of SCN/SDS Resistance in a Plant

10   In accordance with the present invention a method of modulating SCN/SDS resistance in a plant is also provided. The method comprises the step of administering to the plant an effective amount of a substance that modulates expression of an SCN/SDS resistance activity-encoding nucleic acid molecule in the plant to thereby modulate SCN/SDS resistance in the plant.

15   Preferably, the substance that modulates expression of an SCN/SDS resistance activity-encoding nucleic acid molecule comprises a ligand for a modulatable transcriptional regulatory sequence of an SCN/SDS resistance activity-encoding nucleic acid molecule identified in accordance with the methods described above. More preferably, the plant is a soybean plant.

20   Particularly, provided chemical entities (e.g. small molecule mimetics) do not naturally occur in any cell of a lower eucaryotic organism such as yeast. More particularly, provided chemical entities do not naturally occur in any cell, whether of a multicellular or a unicellular organism. Even more particularly, the provided chemical entity is not a naturally occurring molecule, e.g. it is a

chemically synthesized entity. Provided chemical entities can be hydrophobic, polycyclic, or both, molecules, and are typically about 500-1,000 daltons in molecular weight.

XV. Method for Providing SCN/SDS Resistance B Transgenic Plants

5 A "transgenic plant" is a plant that has been genetically modified to contain and express heterologous DNA sequences, either as regulatory RNA molecules or as proteins. As specifically exemplified herein, a transgenic plant is genetically modified to contain and express at least one heterologous DNA sequence operably linked to and under the regulatory control of transcriptional

10 control sequences which function in plant cells or tissue or in whole plants. As used herein, a transgenic plant also refers to progeny of the initial transgenic plant where those progeny contain and are capable of expressing the heterologous coding sequence under the regulatory control of the plant-expressible transcription control sequences described herein. Seeds

15 containing transgenic embryos are encompassed within this definition as are cuttings and other plant materials for vegetative propagation of a transgenic plant.

When plant expression of a heterologous gene or coding sequence of interest is desired, that coding sequence is operably linked in the sense

20 orientation to a suitable promoter and advantageously under the regulatory control of DNA sequences which quantitatively regulate transcription of a downstream sequence in plant cells or tissue or in planta, in the same orientation as the promoter, so that a sense (i.e., functional for translational expression) mRNA is produced. A transcription termination signal, for example,

as polyadenylation signal, functional in a plant cell is advantageously placed downstream of the SCN/SDS resistance coding sequence, and a selectable marker which can be expressed in a plant, can be covalently linked to the inducible expression unit so that after this DNA molecule is introduced into a

5   plant cell or tissue, its presence can be selected and plant cells or tissue not so transformed will be killed or prevented from growing.

In the present invention, the SCN/SDS resistance coding sequence can optionally serve as a selectable marker for transformation of plant cells or tissue.   Where constitutive gene expression is desired, suitable

10   plant-expressible promoters include a native promoter (e.g. SEQ ID NO:15) of the SCN/SDS coding sequences set forth herein as the native promoter is activated in the presence of SCN; the 35S or 19S promoters of Cauliflower Mosaic Virus; the nos, ocs or mas promoters of *Agrobacterium tumefaciens* Ti plasmids; and others known to the art.

15   Indeed, a native promoter (e.g. SEQ ID NO:15) of the SCN/SDS coding sequences set forth herein is activated in the presence of SCN and thus can be used to produce transgenic plants in accordance with the techniques disclosed herein.   Particularly, the native promoter can be linked to a nucleic acid encoding a polypeptide of interest in a construct, and the construct can be used

20   to a prepare a transgenic plant in accordance with techniques described herein. Other techniques are disclosed in U.S. Patent Nos. 5,994,526 and 5,994,527, herein incorporated by reference in their entirety. The polypeptide of interest is then expressed in the plant when the promoter is activated, such as in the presence of SCN or other environmental stimulus.

Where tissue-specific expression of the SCN/SDS resistance coding sequence is desired, the skilled artisan will choose from a number of well-known sequences to mediate that form of gene expression as disclosed herein. Environmentally regulated promoters are also well known in the art, and the

5   skilled artisan can choose from well known transcription regulatory sequences to achieve the desired result.

A method for providing a resistance characteristic to a plant is therefore disclosed. The method comprises introducing to said plant a construct comprising a nucleic acid sequence encoding an SCN/SDS resistance gene

10   product operatively linked to a promoter, wherein production of the SCN/SDS resistance gene product in the plant provides a resistance characteristic to the plant. The construct can further comprises a vector selected from the group consisting of a plasmid vector or a viral vector. The SCN/SDS resistance gene product comprises a protein having an amino acid sequence as set forth as

15   SEQ ID NO:14. The nucleic acid sequence can be a nucleic acid sequence set forth as SEQ ID NO:13, or a nucleic acid that is substantially similar to SEQ ID NO:13, and which encodes an SCN/SDS resistance polypeptide.

The resistance characteristic is preferably nematode resistance, fungal resistance or combinations thereof. More preferably, the nematode resistance

20   is *H. glycines* resistance or root knot nematode resistance.

In an alternative embodiment, the construct further comprises another nucleic acid molecule encoding a polypeptide that provides an additional desired characteristic to the plant. Other desired characteristics include yield, drought resistance, chemical resistance (e.g. herbicide or pesticide resistance),

spoilage resistance or any or other desired characteristic as would be apparent to one of ordinary skill in the art after review of the disclosure of the present invention. Representative nucleic acids sequences are described in the following U.S. patents (incorporated herein by reference in their entirety): U.S.

5   Patent No. 5,948,953 to Webb (brown rot fungus resistance); U.S. Patent No. RE36,449 to Lebrun et al. (herbicide resistance); U.S. Patent No. 5,952,546 to Bedbrook et al. (delayed ripening tomato plants); and U.S. Patent No. 5,986,173 to Smeekens et al. (transgenic plants showing a modified fructan pattern).

10   Optionally, the method further comprises monitoring an insertion point for the construct in the plant genome; and providing for insertion of the construct into the plant genome at a location not associated with the resistance characteristic, the desired characteristic, or both the resistance or the desired characteristic.

15   XVI.   Method for Providing SCN/SDS Resistance B Marker-Assisted Selection and Development of a Breeding Program

The present invention relates to a novel and useful method for introgressing, in a reliable and predictable manner, SCN/SDS resistance into non-resistant soybean germplasm. The method involves the genetic mapping of

20   loci associated with SCN/SDS resistance, definition of genetic markers that are linked with SCN/SDS resistance, and a high-throughput PCR-based assay for detecting such a genetic marker. Markers useful in a preferred embodiment of the invention include the following: a locus mapping to linkage group G and mapped by one or more of the markers set forth SEQ ID NOs:1-6, a locus

mapping to linkage group A2 and mapped by one or more of the markers set forth as SEQ ID NOs:7-12; or combinations thereof. Also preferably, a genetic marker used for marker-assisted selection comprises a sequence, or portion thereof, of any one of SEQ ID NOs:13 and 16-19, or combinations thereof.

5   From the sequence data found in SEQ ID NOs:1-13 and 16-19, and from the other markers identified herein, primer pairs, as for example, PCR primer pairs, capable of distinguishing differences among these genotypes are developed. Simple assays for the markers and genes use a label, such as, but not limited to, a covalently attached chromophores, that do not need

10   electrophoresis are developed to increase the capacity of marker assisted selection to help plant breeders. A preferred assay is the TaqMan™ assay disclosed in Example 6. Non-destructive sampling of dried seed for DNA preparations are developed to allow selection prior to planting, for example, using the methods set forth in Example 9. This enables the testing of the

15   effectiveness of marker assisted selection in predicting field resistance to SCN and SDS.

      A preferred manner for providing SCN/SDS resistance to a plant involves providing one or more plants from a parental soybean plant line which comprises in its genome one or more molecular markers comprising a

20   sequence, or portion thereof, set forth as any one of SEQ ID NOs:1-13 and 16-19. Preferably, the parental plant is purebreeding for one or more of the molecular markers, more preferably the parent plant is purebreeding for molecular markers comprising a sequence, or portion thereof, set forth as any

one of SEQ ID NOs:1-13 and 16-19. In one preferred embodiment, the parental line is "Forrest" or a line derived therefrom.

The SCN/SDS resistance trait can be introgressed into a recipient soybean plant line which is non-resistant or less resistant to SCN/SDS by performing marker-assisted selection based on the molecular markers of the present invention as set forth as SEQ ID NOs:1-13 and 16-19.

Introgressing can be accomplished by any method known in the art, including but not limited to single seed descent, pedigree method, or backcrossing, each described herein below. Additional methods for introgressing are disclosed in U.S. Patent Nos.5,948,953 and 6,162,967. Any suitable method can be used, the critical feature being marker-assisted selection of a marker of the present invention using a nucleotide sequence assay.

Single Seed Descent. According to this method, "Forrest" can be crossed to "Essex", and the seed planted in a field. The resulting seed (F2) is planted in the greenhouse and the resulting seeds (F3) are harvested while keeping separate the seeds from each plant. A random F3 seed from each of approximately 200 plants is planted and the resulting F4 seed is harvested. The seeds from each individual plant are again kept separate. A random F4 seed from each of the approximately 200 plants is planted and the resulting F5 seed is harvested. This selection process is repeated until F7 seed is harvested and identified as an inbred line. At each generation beginning with the F3 generation, plants are screened with soybean cyst nematodes, and plants were selected for advancement based upon the presence of SCN resistance and

other phenotypic characteristics. Alternatively, plants are screened for the presence of one or more of the molecular markers listed herein using a TaqMan™ genotyping assay and selected for advancement based upon the presence of one or more of the markers.

5   Pedigree Method. Using a SCN resistant recombinant inbred line, produced for example by single seed descent, as a donor source, the SCN resistant trait can be introgressed into other germ plasm sources. To develop new germplasm, the SCN resistant recombinant inbred line is used as one of the parents. The resulting progenies are evaluated and selected at various

10   locations for a variety of traits, including SCN resistance. SCN resistance is determined by phenotypic screening or by genotyping based upon the presence of the molecular markers listed herein.

      Backcrossing. Using a SCN resistant recombinant inbred line, produced for example by single seed descent, as a donor source, the SCN resistant trait

15   is introgressed into other soybean plant lines. The SCN resistant recombinant inbred line is crossed to a line that demonstrates little or non SCN resistance (the recipient). The resulting plants are crossed back to the recipient soybean plant line that is being converted to SCN resistance. This crossing back to the parental line that is being converted may be repeated several times. After each

20   round of backcrossing, plants are selected for SCN resistance, which can be determined by either phenotypic screening or by the selection of molecular markers linked to SCN resistance loci. Besides selecting for SCN resistance, the plants are also selected that most closely resemble the original plant line

being converted to SCN resistance. This selection for the original plant line is done phenotypically or with molecular markers.

In one specific preferred method, BC_NF1 plants are genotypically screened for the presence of one or more markers linked to SCN resistance

5   genomic loci. As used herein, the term "BC_NF1 plant" is intended to refer to a plant in the first generation after a specific backcross event, the specific backcross event being designated by the term "N", irrespective of the number of previous backcross events employed to produce the plant. Plants having the one or more markers present may preferably be backcrossed with plants of the

10   parental line or, alternatively, be selfed, the plants resulting from either of these events also being genotypically screened for the presence of one or more markers linked to SCN resistance genomic loci. This procedure can be repeated several times.

In another specific preferred method, BC_NF1 plants are selfed to produce

15   BC_NF2 seeds. BC_NF2 plants are then screened either genotypically using, for example a TaqMan^{TM} assay as disclosed in Example 6, or by phenotypic assessment of SCN resistance. Those plants having present one or more molecular markers linked to SCN resistance, or those plants displaying resistance, depending upon the screening method used, are backcrossed with

20   plants of the parental line to produce BC_NF3 seeds and plants. This procedure can be repeated several times. In a soybean breeding program, the methods of the present invention can be used for marker-assisted selection of the molecular markers described herein. Genetic markers closely linked to SCN/SDS resistance genes can be used to indirectly select for favorable alleles more

efficiently than phenotypic selection. Genetic markers comprising SCN/SDS resistance genes, as disclosed herein, can be used to select for SCN/SDS resistance genes with optimal efficiency and accuracy.

Marker-assisted selection can be employed to select one or more loci at a wide variety of population development stages in a two-parent population, multiple parent population, or a backcross population. Such populations are described in Fehr (1987) Breeding Methods for Cultivar Development J. R. Wilcox (ed.) and Soybeans: Improvement, Production, and Uses, 2nd ed..

Marker-assisted selection according to art-recognized methods can be made, for example, step-wise, whereby the different SCN resistance loci are selected in more than one generation; or, as an alternative example, simultaneously, whereby all loci are selected in the same generation. Marker-assisted selection for SCN resistance can be done before, in conjunction with, or after testing and selection for other traits such as seed yield, plant height, seed type, etc. The DNA from target populations, isolated for use in accordance with genetic marker detection, can be obtained from any plant part, and each DNA sample can represent the genotype of single or multiple plant individuals, including seed.

Marker-assisted selection can also be used to confirm previous selection for SCN resistance or susceptibility made by challenging plants with SCNs in the field or greenhouse and scoring the resulting phenotypes. Alternatively, plants can be analyzed by TaqMan™ genotyping to determine the presence of the above-described molecular markers, thus confirming the presence of a genomic locus associated with SCN resistance.

As such, also provided by the present invention are methods for determining the presence or absence of SCN resistance in a soybean plant, or alternatively in a soybean seed. These methods comprise analyzing genomic DNA from a plant or a seed for the presence of one or more of the molecular

5   markers set forth as SEQ ID NOs:1-13 and 16-19. According to this method, the analyzing comprises performing a TaqMan™ assay as disclosed in Example 6, or any other suitable method known in the art.

The ability to distinguish heterozygotes and their derived heterogeneous lines is important to early generation selection (before the F_5) in soybean

10   breeding programs when within population variability is high (Bernard et al. (1988) *USDA Tech Bull* 1796; Brown et al., 1987). The lower stringency TaqMan™ 2 assay disclosed herein was most effective for identifying most of the heterogeneous lines in this population. However, the cutoff values of FAM and TET for the efficient identification of heterogeneous lines (or heterozygous

15   F2 lines) is likely to vary across assays and should be set arbitrarily according to expectations of the number of lines that are expected to contain both alleles. The assay was used for analyzing 2,000 lines derived from specific cultivar crosses over 3 days. A single researcher can process 768 sample per day (8 x 96 samples) since the reading time of the machine is 15 minutes for one 96 well

20   plate and the thermal cycler stage takes about 2 hours.

Table 3 shows that with genomic DNA from 94 cultivars the standard TaqMan™ allelic discrimination assays and PCR assays provided allele scores that were in good agreement with the cultivar phenotypes (Concibidio, 1997; Bernard et al., 1988). Cultivars, plant introductions (PI), breeding lines and

germplasm releases listed in Table 3 were parents in the SCN molecular breeding program at Southern Illinois University-Carbondale (SIUC) from 1997-1999. The prevalence of allele 1 was in good agreement with allele frequencies for markers that are closely linked to *Rhg4* (Cregan et al. 1999; Mathews et al.

5 (1998) *Theor Appl Genet* 97:1047-1052; Mahalingam et al., 1995). Those resistant cultivars sharing allele 1 with the susceptible lines may not require the presence of *Rhg4* for resistance to SCN or have derived their resistance to SCN at the *Rhg4* locus from alleles derived from cultivars other than Forrest. In addition, some soybean breeders may have been effective in separating even

10 the most closely linked marker from resistance genes using phenotypic selection. However, this is probably infrequent since selection to generate the resistance allele 2 in susceptible cultivars has not occurred frequently. Only three cultivars with allele 2 were susceptible.

Table 3

	Resistant	Susceptible
Allele 2	Forrest, Hartwig, Fayette, Pharaoh, Picket, Accomac, Bedford, Delsoy4710, Peking, PI88788, PI209332, PI90763, PI437654, LS92-1088 , LS92-4173, LS94-3207, LS95-0259, LS95-0709, LS95-1454, LS96-1631, LS90-1920, LS94-3545, S92-1679, S92-2711A, S94-2086, LN94-10527, A5560K1390, K1425	MD93-5298 Pace Holladay
Allele 1	Manokin, Mustang, Dwight, Pana, Ina, PI 398680, IA2036, IA3005, LS92-3660, LS93-0292, LS93-0375, LS94-2435, LS96-0735, LS96-3813, LS96-5009, LN92-10725, GX93-1573, SS94-7546, SS94-4337, S95-1908, A4138, A95-483010, M92-1645, M92-1708, M90-184111, K1423, K1424	Essex, Bragg, Dunfield, Hill, CNS, Lee, Noir1, Ogden, Calhoun, Chesapeake, Choska, Stressland, Macon, Misuzudaiza, Nakasennari, PI 520733, PI567445B, PI567583C, PI567650B, PI 567374, PI 567650B, IA3010, IA1006, TN96-58, N96-180, LN93-11632, LN93-11945, LN95-5417, A94-674017, A94-774021, A96-494018, C1963, HC93-2690,HS93-4118, K1410

Summarily, the sequences and methods disclosed herein enable automated, high throughput, rapid genotyping of DNA polymorphisms for selection of SCN/SDS resistance in breeding programs.

5   Examples

The following Examples have been included to illustrate preferred modes of the invention. Certain aspects of the following Examples are described in terms of techniques and procedures found or contemplated by the present inventors to work well in the practice of the invention. These Examples are

10   exemplified through the use of standard laboratory practices of the inventors. In light of the present disclosure and the general level of skill in the art, those of skill will appreciate that the following Examples are intended to be exemplary only and that numerous changes, modifications and alterations can be employed without departing from the spirit and scope of the invention.

15   EXAMPLE 1

Plant Material

A mapping population consisted of approximately 100 recombinant inbred lines derived at the F5 generation from a cross of 'Essex' (Smith & Camper (1973) Crop Sci 13:459) by 'Forrest' (Hartwig & Epps (1973) Crop Sci 13:287).

20   The recombinant inbred line (RILs) population was advanced to the F5:13 generation from 300 plants per RIL per generation (Hnetkovsky et al., 1996). Forrest is resistant to the soybean cyst nematode (SCN) populations classified as race 3 and Essex is susceptible to all populations of SCN (Chang et al., 1997; Meksem et al. 1999).

EXAMPLE 2

SCN Female Index (FI) Determination

The number of white female cysts was compared on each genotype to the number of white female cysts on a susceptible control, such as Essex, to

5   determine the female index (FI) for each population (Meksem et al., 1999). Seedlings were inoculated with 2000+/-25 eggs from a homogenous isolate of *H. glycines*. All experiments used five single-plant replications per line. The mean number of white female cysts on each genotype and the susceptible control were determined and FI was calculated as the ratio of the mean number

10   of cysts on each genotype to the mean number of cysts on the susceptible check.

EXAMPLE 3

Characterization of New Markers for SCN/SDS Resistance

Soybean genomic DNA used for AFLP analysis was extracted and

15   purified using the Qiagen (Hilden, Germany) Plant Easy DNA Extraction Kit. Primary template DNA was prepared using the restriction enzymes *Eco*RI and *Mse*I.

AFLP analysis was performed as described by Vos et al. (1995) *Nuc Acids Res* 23:4407-4414 except that the streptavidin bead selection step was

20   omitted. PCR reactions were performed with using primer pairs derived from each of two sets of primers. Primers within *Eco*RI set all included the core sequence E: 5'- GAC TGC GTA CCA ATT C (SEQ ID NO:115) with 1 or 3 base pair extensions. Primers of the *Mse*I set have the sequence M: 5'- GAT GAG TCC TGA GTA A (SEQ ID NO:116) with 1 or 3 base pair extensions. The

primer combinations (EA and MC) and (EC and MA) were used for pre-amplification of primary template. Three selective nucleotides per primer were used to generate AFLP fragments from the secondary templates. AFLP bands were labeled with ^{33}P by primer phosphorylation, separated by

5   electrophoresis on 4% (w/v) PAGE and visualized by exposing X-ray film to the dried gel.

Target AFLP bands on the autoradiograph were matched to the corresponding area in the gel and the appropriate AFLP fragment was excised from the dried gel. The band was eluted from the gel by incubation in 100ml of

10   water at 4°C for 1 hour. Sequence isolation in bacterial clones was performed as described by Meksem et al. (1995) *Mol Gen Genet* 249:74-81 with the modification that the pGEM-T vector (Promega, Madison, Wisconsin) was ligated to PCR amplified, gel eluted DNA. DNA sequencing of clones allowed PCR primers to be designed for each unique DNA sequence using Oligo 5.0

15   software (PE Biosystems, Foster City, California). The PCR product was analyzed on 4% (w/v) Metaphor7 (FMC, Rockland, Maine) agarose gel.

AFLP markers that were dominant or co-dominant, in repulsion and in coupling phases were used. For dominant AFLP markers, the band of the dominant allele was cloned and sequenced. The corresponding marker for the

20   recessive allele was isolated by PCR using primers designed from the dominant band sequence. For apparently co-dominant AFLP markers, both, the coupling and repulsion phase bands were cloned simultaneously from the acrylamide gel.

The general strategy employed to identify the specific sequence underlying AFLP band polymorphisms was as follows. If the polymorphism was

dominant (e.g. $E_{ATG}M_{CGA}87$) a primer pair was designed to flank each of the unique sequences derived from the AFLP band. Each primer pair was used to amplify genomic DNA from both Essex and Forrest. Any primer set that revealed polymorphism (dominant or co-dominant) between the two parents was used to amplify members of the RIL mapping population. The primer pair that generated a marker on the map corresponding to the map position of the original AFLP band was inferred to be the specific marker STS.

For some AFLP bands the above strategy was ineffective, presumably because polymorphism was within or close to the restriction site used for AFLP linker ligation (e.g. $E_{CGG}M_{AGA}116$). In such cases genomic DNA from the parents and mapping population was used in a modified AFLP protocol as follows. The pre-amplification step was omitted and the six selective nucleotide step was replaced by an extended highly selective MseI primer to which we added the first 7 bases of the sequenced band, combined with a non selective EcoRI primer E (e.g. MseI primer M AGAGACT and EcoRI primer E). The MseI primer was end-labeled by phosphorylating the 5' end with 5 ml [g-^{33}P] ATP (3000 Ci/mmol) for 30 min at 37 °C with 10 units of T4 Kinase (Pharmacia, Piscataway, New Jersey). Any primer set that revealed polymorphism (dominant or co-dominant) between the two parents was used to amplify members of the RIL mapping population. The primer pair that generated a marker on the map corresponding to the map position of the original AFLP band was inferred to be the specific marker STS.

EXAMPLE 4

Cloning of SCN/SDS Resistance Genes in Linkage Groups G and A2

The cloned AFLP bands of Example 3 were used to screen the soybean

Forrest *Bam*HI or *Hind*III BAC libraries by PCR as described by Meksem et al.

5   (2000).

Both plasmid and BAC DNA was prepared using the appropriate kit

(Qiagen, Hilden, Germany). Sequence determinations were performed by the

di-deoxy chain-termination method using Advanced Biosystems (ABI, Foster

city, California) "big dye" cycle sequencing separated on ABI 377 automated

10   DNA sequencer.

Plasmids containing clones derived from AFLP bands were sequenced

using M13 universal forward and reverse primers. Direct BAC insert sequencing

was performed as above with the following modifications: BAC DNA was heated

for 30 min at 70 °C, and sheared by pippeting into a narrow gauge tip for 2 min.

15   Two primers designed from the target AFLP band sequence were used for

sequencing. For the $E_{ATG}M_{CGA}87$ positive BAC insert DNA, the forward primer,

named ATG4BACF (SEQ ID NO:117), was 5' gggtttcagataaccgtggtcg 3', the

reverse primer was the complementary strand sequences of the ATG4BACF

primer. The PCR conditions used was 95°C for 10 min, then 45 cycles of 95°C

20   for 30 sec, 55°C for 20 sec and 60 °C for 4 min.

EXAMPLE 5

Taqman™ Genotyping Assay

PCR primers and TaqManTM probes were designed with the primer

express program (Perkin-Elmer/Applied Biosystems, Foster City, California) and

were custom synthesized by Perkin-Elmer. Two TaqMan™ probes were designed to encompass the A2D8 (Figure 1) insertion polymorphisms (underlined). The A2D8 SCAR was derived from the codominant AFLP bands E$_{CCG}$-M$_{AAC}$417 (Essex, allele 1, GenBank Accession No. AF286701) and

5   E$_{CCG}$-M$_{AAC}$409 (Forrest, allele 2, GenBank Accession No. AF286700) that contain a homolog (P= 2e-05) of one component (Tic22; GenBank Accession No. AAC64606.1) of the protein import apparatus of the chloroplast inner envelope membrane. Allele 1: 5'-TET-TTG CAG ATA TTT TAG TTG ATT GGC C-TAMRA (SEQ ID NO:118). Allele 2: 5'-6FAM-AGT TGA TTG GCT CAA ACC

10   ATG GCC-TAMRA (SEQ ID NO:119). Reverse Primer:' 5' d TTG CGT GTG ATC GGT ATT AC 3' (SEQ ID NO:120). Forward primer: 5' d T ACC TGA GTT CTC TCA AGT C 3' (SEQ ID NO:121).

TaqMan™ reactions were performed essentially as the Perkin-Elmer TaqMan™ PCR Reagent Kit protocol describes except the PCR reaction was

15   performed in 384 well plates to reduce assay volume and cost. Briefly, each reaction contained 10ng of the extracted DNA, 0.025 units/ml of AmpliTaq Gold™ (Perkin-Elmer/Applied Biosystems, Foster City, California), 400nM of the forward and reverse primers (Research Genetics, Huntsville, Alabama), 50nM of FAM fluorescent probe and 150 nM of TET fluorescent probe

20   (Perkin-Elmer/Applied Biosystems, Foster City, California) in 1 x universal master mix (Perkin-Elmer/Applied Biosystems, Foster City, California). The above ratio of primers and probes was optimized using a series of primer/probe combinations to reach a maximal signal and the balance of the two probes by reading in an ABI 7200 sequence detector. The TaqMan™ universal PCR

master mix is a premix of all the components, except primer and probes, necessary to perform a 5' nuclease assay. The final optimized conditions represented a two step PCR protocol, with two holds followed by cycling, on a 384 well thermal cycler (GeneAmp PCR System 9700, Perkin-Elmer/Applied

5   Biosystems, Foster City, California). The two hold cycles were 50°C for 2 min and 95°C for 10 min. The 35 cycles were at 95°C for 15 sec, 60°C for 1min. After amplification the plates were cooled to room temperature and samples were transferred from a 384 well plate to a 96 well MicroAmpJ optical tray and fluorescence was detected on an ABI PrismJ 7200 Sequence Detector

10   (Perkin-Elmer/Applied Biosystems, Foster City, California).

The results were analyzed by allelic discrimination of the sequence detection software (Perkin-Elmer/Applied Biosystems, Foster City, California). Two grouping methods were used to attempt to accurately separate heterogeneous lines from homogeneous lines at each allele. In grouping

15   method 1 (TaqMan^{TM} 1) a stringent cut-off for FAM (>7) was used for allele 1 compared to heterogenous scores. This served to reduce the number called as potentially heterogeneous to about the percentage expected from the breeding method used for RIL development (6%). Fluorophore ratios were as follows; no amplification (FAM and TET both less than 6 units); allele 1 homozygous (FAM

20   less than 7, TET greater than 7); allele 2 homozygous (FAM greater than 10, TET less than 5); and heterogeneous for allele 1 and allele 2 (FAM greater than 7, TET 5-8). For TaqMan^{TM} selection grouping method 2 ratios were; no amplification (FAM and TET both less than 6 units); allele 1 homozygous (FAM less than 5, TET greater than 7); allele 2 homozygous (FAM greater than10,

TET less than 5); and heterogeneous for allele 1 and allele 2 (FAM greater than 5, TET 5-9). The FAM and TET signals were stable in the dark for 2 days after PCR.

EXAMPLE 6

5   Genotyping Assay Using Gel Electrophoresis Markers

PCR reactions were performed with DNA from the recombinant inbred lines. The 114 and 120 base pair PCR products were generated using the forward and reverse primers (SEQ ID NOs:120-121). The final optimized conditions were 94°C for 10 min, then 35 cycles of 94°C for 25 sec, 56°C for 30 10   sec and 72 °C for 60 sec. After the PCR reactions were completed, the plates were cooled to room temperature and the PCR products separated by electrophoresis on a 4% (w/v) agarose gel.

EXAMPLE 7

Allele Distribution in Soybean Germplasm

15   Genotypes at A2D8 were determined from the genomic DNA of 94 cultivars that represented the parents of populations in the SIUC soybean breeding program from 1997-1999 (Table 3). There were 38 cultivars susceptible to SCN and 56 cultivars resistant to SCN race 3. Allele 2 (R) was found in 32 of 94 cultivars tested. There were very few susceptible genotypes 20   with allele 2 (3 of 32) and the majority of genotypes with allele 2 (29 of 32) were resistant to SCN. In contrast, allele 1 (S) was found in 62 cultivars but frequently in both resistant cultivars (27 of 56) and susceptible cultivars (35 of 38).

EXAMPLE 8

Selection of SCN/SDS Resistant Seeds

G.max L. seeds used to start cultures should be less than six months old and have been stored in darkness at 4°C. Then, the seeds are cultured as

5   folllows:

1. Surface disinfect with 70% (v/v) ethanol for 2 min then 20% (v/v) bleach for 20 min. Rinse three times in sterile MS media.

2. Germinate the seed on MS media containing 10 g/l agar, 30 g/l sucrose but no PGRs for 3 days at 27°C.

10   3. Axenically remove the testa, remove the cotyledonary notes, cut the cotyledons transversely in half and use the distal cotyledonary halves to establish callus cultures.

To initiate callus growth, cotyledonary halves are placed on MS medium with 30 g/l sucrose, 5 mM kinetin, 100 mg/l myoinositol, 0.5 mg/mL thiamine•HCl

15   pH 5.7 at 27°C unless noted below. The medium contains 5 mM indolebutyric acid as auxin. Place cotyledonary halves in tubes containing 10 mL solidified media. Incubate for 28 days.

To assay callus growth, pieces of callus each approximately 25 mg should be added to sterile tubes containing 10 mL media with varying

20   concentrations of *H. glycines*, *F.solani* or extracts thereof. After 28 days at 28°C the explants are evaluated for growth and growing sectors subcultured.

Cell suspensions are derived by placing 2 g of a macerated callus in 40 mL of MS medium. The flask, a 125 mL Erlenmeyer flask, should be capped with a foam plug. Subcultures should be made every 14 days into fresh media by allowing the cells to settle, removing the old media by aspiration, adding

5   twice the volume of fresh media and splitting into two flasks.

Soybean tissue capable of regeneration to whole plants are grown in the presence of *H. glycines*, *F.solani* or extracts thereof. Cell lines representing mutants capable of continued growth are regenerated and the heritability of SCN or SDS resistance determined in these plants or their seed or tissue

10   derived progeny.

REFERENCES

The publications and other materials listed below and/or set forth in the text above to illuminate the background of the invention, and in particular cases,

15   to provide additional details respecting the practice, are incorporated in their entirety herein by reference. Materials used herein include but are not limited to the following listed references.

Adelman et al. (1983) *DNA* 2:183-193.

Alam and Cook (1990) *Anal Biochem* 188:245-254.

20   Altschul et al. (1990) *J Mol Biol* 215:403-410.

Arondel et al. (1992) *Science* 258:1353-1355.

Ausubel et al. (1992) Current Protocols in Molecular Biology, John Wylie and Sons, Inc., New York.

Bartlett et al. (1982) in Methods in Chloroplast Molecular Biology, Edelmann et al. (Eds.), pp 1081-1091, Elsevier.

Barton (1998) *Acta Crystallogr D Biol Crystallogr* 54:1139-1146.

Batzer et al. (1991) *Nucleic Acid Res* 19:3619-3623.

Bell-Johnson et al. (1998) *Soybean Genet Newslett* 25:115-118.

Bernard et al. (1988) *USDA Tech Bull* 1796.

Bevan (1984) *Nucl Acids Res* 12:8711-8721.

Bevan et al. (1983) *Nature* 304:184-187.

Binet et al. (1991) *Plant Science* 79: 87-94.

Blochlinger & Diggelmann (1984) *Mol Cell Biol* 4:2929-2931.

Bourouis et al., (1983) *EMBO J* 2(7):1099-1104.

Bodanszky, et al. (1976) Peptide Synthesis, John Wiley and Sons, Second Edition, New York.

Brookes (1999) *Gene* 234(2):177-186.

Brown et al. (1987) Principles and Practice of Nematode Control in Crops, pp179-232, Academic Press, Orlando Florida.

Callis et al. (1987) *Genes Develop* 1:1183-1200.

Chang et al. (1996) *Crop Sci* 36:965-971.

Chang et al. (1997) *Crop Science* 37(3):965-971.

Chase et al. (1997) *Theor Appl Genet* 94:724-730.

Chibbar et al. (1993) *Plant Cell Rep* 12:506-509.

Christou et al. (1991) *Biotechnology* 9:957-962.

Comai et al. (1988) *J Biol Chem* 263:15104-15109.

Concibido (1996) *Theor Appl Genet* 93:234-241.

Concibido et al. (1997) *Crop Sci* 37:258-264.

Conner et al. (1983) *Proc Natl Acad Sci USA* 80:278-282.

Cregan et al. (1999a) *Crop Sci* 39:1464-1490.

Cregan et al. (1999b) *Theor Appl Genet* 99:811-818.

5   Cregan et al. (1999c) *Theor Appl Genet* 99:918-928.

Cubitt et al. (1995) *Trends Biochem Sci* 20:448-455.

Datta et al. (1990) *Biotechnology* 8:736-740.

EP 0 292 435

EP 0 332 104

10   EP 0 332 581

EP 0 342 296

EP 0 392 225

EP 0 452 269

Fehr (1987) in Soybeans: Improvement Production and Uses 2nd Ed.,

15   J.R. Wilcox (ed.), American Society of Agronomy, Madison, Wisconsin.

Firek et al. (1993) *Plant Molec Biol* 22:129-142.

Fromm et al. (1990) *Biotechnology* 8:833-839.

Gallie et al. *(1987) Nucl Acids Res* 15:8693-8711.

Glover, ed. (1985) DNA Cloning: A Practical Approach, MRL Press, Ltd.,

20   Oxford, United Kingdom.

Gomez A.K. and A.A. Gomez (1984) *Statistical Procedures For Agricultural Research.* 2nd ed. John Wiley & Sons New York

Gordon-Kamm et al. (1990) *Plant Cell* 2:603-618.

Gritz et al. (1983) *Gene* 25:179-188.

Harlow and Lane (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Hartwig and Epps (1973) *Crop Science* 13:287.

Henikoff et al. (2000) *Electrophoresis* 21(9):1700-1706.

5 Henikoff and Henikoff (1989) *Proc Natl Acad Sci USA* 89:10915.

Henikoff and Henikoff (2000) *Adv Protein Chem* 54:73-97.

Hnetkovsky et al. (1996) *Crop Science* 36(2):393-400.

Hofgen & Willmitzer (1988) *Nucl Acids Res* 16:9877.

Huang et al. (2000) *Pac Symp Biocomput* 230-241.

10 Hudspeth & Grula (1989) *Plant Molec Biol* 12:579-589.

Hutchens and Yip (1993) *Rapid Commun Mass Spectrom* 7: 576-580.

Kalinina et al. (1997) *Nucl Acids Res* 25:1999-2004.

Kanazin et al. (1996) *Proc Natl Acad Sci USA* 93(21):11746-11750.

Karlin and Altschul (1993) *Proc Natl Acad Sci USA* 90:5873-87.

15 Keim et al. (1997) *Crop Science* 37:537-543.

Kestila et al. (1998) *Mol Cell* 1(4):575-582.

Klein et al. (1987) *Nature* 327:70-73.

Koduri and Poola (2001) *Steroids* 66(1):17-23.

Koziel et al. (1993) *Biotechnology* 11:194-200.

20 Kyte et al. (1982) *J Mol Biol* 157:105.

Landers & Botstein (1989) *Genetics* 121:185-199.

Landgren et al. (1988) *Science* 241:1007.

Landgren et al. (1988) *Science* 242:229-237.

Landegren et al. (1998) *Genome Res* 8:769-776.

Lark et al. (1993) *Theor Appl Genet* 86:901-906.

Li and Herskowitz (1993) *Science* 262:1870-1874.

Liedberg et al. (1983) *Sensors Actuators* 4:299-304.

Livak et al. (1995) *PCR Meth and Applic* 4:357-362.

5   Livak et al. (1995) *Nat Genet* 9:341-342.

Logemann et al. (1989) *Plant Cell* 1:151-158.

Luo et al. (1999) *Plant Disease* 83:1155-1159.

Madge et al. (1972) *Phys Rev Lett* 29:705-708.

Mahalingam et al. (1995) *Breed Sci* 45:435-445.

10   Mahalingham et al. (1996) *Genome* 39:986-998

Maiti et al. (1997) *Proc Natl Acad Sci USA*, 94:11753-11757.

Malmquist (1993) *Nature* 361:186-187.

Martin et al. (1993) *Science* 262:1432-1436.

Mathews et al. (1998) *Theor Appl Genet* 97:1047-1052.

15   Matthews et al. (1991) *Soybean Genetics Newsletter.*

McBride et al. (1990) *Plant Molecular Biology* 14:266-276.

McElroy et al. (1990) *Plant Cell* 2:163-171.

McElroy et al. (1991) *Mol Gen Genet* 231:150-160.

Meksem et al. (1995) *Mol Gen Genet* 249:74-81.

20   Meksem et al. (1999) *Theor Appl Genet* 99:1131-1142.

Meksem et al. (2000) *Theor Appl Genet* 101:747-755.

Messing & Vierra (1982) *Gene* 19:259-268.

Myers & Anand (1991), *Euphytica* 55:197-201.

Nasarabadi et al. (1999) *BioTechniques* 27:1116-1117.

Needleman & Wunsch (1970) *J Mol Biol* 48:443-453.

Njiti et al. (1996) *Crop Science* 36:1165-1170.

Ochman et al. (1990) in <u>PCR protocols: a Guide to Methods and Applications</u>, Innis et al. (eds.), pp. 219-227, Academic Press, San Diego, California.

Ohtsuka et al. (1985) *J Biol Chem* 260:2605-2608.

Orita et al. (1989) *Proc Natl Acad Sci USA* 86(8):2766-2770.

Paszkowski et al. (1984) *EMBO J* 3:2717-2722.

Paterson et al. (1990) *Genetics* 124:735-742.

Pearson & Lipman (1988) *Proc Natl Acad Sci USA* 85:24442448.

Potrykus et al. (1985) *Mol Gen Genet* 199:169-177.

Prabhu et al. (1999) *Crop Science* 39(4):982-987.

Price (1993) *Blood Rev* 7:127-134.

Rao-Arrelli et al. (1988) *Crop Science* 28:650-652.

Rao-Arrelli et al. (1992) *Crop Science* 32:862-864.

Regan et al. (2000) *Anal Biochem* 286(2):265-276.

Reich et al. (1986) *Biotechnology* 4:1001-1004.

Riggs and Schmidt (1988) *J Nematol* 20:392-395.

Rogers et al. (1989) *Proc Natl Acad Sci USA* 82:6512-6516.

Rohrmeier & Lehle (1993) *Plant Molec Biol* 22:783-792.

Rommens et al. (1989) *Science* 245:1059-1065.

Rose & Botstein (1983) *Meth Enzymol* 101:167-180.

Rossolini et al. (1994) *Mol Cell Probes* 8:91-98.

Rothstein et al. (1987) *Gene* 53:153-161.

Saiki et al. (1985) *Bio/Technology* 3:1008-1012.

Sambrook et al. eds. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, New York.

Saqi et al. (1999) *Bioinformatics* 15:521-522.

Sauer (1998) *Methods* 14(4):381-392.

Schmidhauser & Helinski (1985) *J Bacteriol* 164:446-455.

Schocher et al. (1986) *Biotechnology* 4:1093-1096.

Shimamoto et al. (1989) *Nature* 338:274-277.

Shinshi et al. (1990) *Plant Molec Biol* 14:357-368.

Shoemaker et al. (1995) *Crop Science* 35:436-446.

Silhavy et al. (1984) Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, New York.

Singh et al. (1989) *Biotechniques* 7:252-261.

Skuzeski et al. (1990) *Plant Molec Biol* 15:65-79.

Smith & Waterman (1981) *Adv Appl Math* 2:482.

Smith & Camper (1973) *Crop Science* 13:459.

Spencer et al. (1990) *Theor Appl Genet* 79:625-631.

Staskawicz (1995) *Science* 268:661-667.

Stoneking et al. (1991) *Am J Hum Genet* 48(2):370-82.

Thompson et al. (1987) *EMBO J* 6:2519-2523.

Tijssen (1993) in Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes, part 1 chapter 2, Elsevier, New York, New York.

Trask (1991) *Trends Genet* 7:149-154.

Uknes et al. (1992) *The Plant Cell* 4:645-656.

Uknes et al. (1993) *The Plant Cell* 5:159-169.

Unger et al. (1989) *Plant Molec Biol* 13:411-418.

5   U.S. Patent No. 4,196,265

U.S. Patent No. 4,554,101

U.S. Patent No. 4,940,935

U.S. Patent No. 4,945,050

U.S. Patent No. 5,036,006

10   U.S. Patent No. 5,100,792

U.S. Patent No. 5,188,642

U.S. Patent No. 5,260,203

U.S. Patent No. 5,523,311

U.S. Patent No. 5,591,616

15   U.S. Patent No. 5,614,395

U.S. Patent No. 5,629,158

U.S. Patent No. 5,639,949

U.S. Patent No. 5,834,228

U.S. Patent No. 5,872,011

20   U.S. Patent No. 5,948,953

U.S. Patent No. 5,952,546

U.S. Patent No. 5,958,624

U.S. Patent No. 5,986,173

U.S. Patent No. 5,994,526

U.S. Patent No. 5,994,527

U.S. Patent No. 6,096,555

U.S. Patent No. 6,162,967

U.S. Patent No. RE36,449

5   van den Broeck et al. (1985) *Nature* 313:358-363.

Vasil et al. (1992) *Biotechnology* 10:667-674.

Vasil et al. (1993) *Biotechnology* 11:1553-1558.

Vidal et al. (1996) *Proc Natl Acad Sci USA* 93(19):10315-10320.

Vos et al. (1995) *Nucleic Acids Research* 23:4407-4414.

10   Wang et al. (1998) *Science* 280(5366):1077-82.

Warner et al. (1993) *Plant J* 3:191-201.

Webb et al. (1995) *Theor Appl Genet* 91:574-581.

Weeks et al. (1993) *Plant Physiol* 102:1077-1084.

Weiseman et al. (1992) *Theor Appl Genet* 85:136-138

15   White et al. (1990) *Nucl Acids Res* 18:1062.

WO 93/07278

WO 93/21335

WO 94/00977

WO 97/47763

20   Worrall et al. (1998) *Anal Biochem* 70:750-756.

Wrather et al. (1995) *Plant Disease* 79:1076-1079.

Xu et al. (1993) *Plant Molec Biol* 22:573-588.

Yuan et al. (1999) *Hum Mutat* 14(5):440-446.

Zhang et al. (1988) *Plant Cell Rep* 7:379-384.

Zhang et al. (1994) *Mol Gen Genet* 244:613-621.

Zhu et al. (1996) *Mol Gen Genet* 252:483-488.

Zimmer et al. (1993) Peptides pp. 393B394, ESCOM Science Publishers, B. V.

5   Zobrist et al. (2000) *Soybean Genet Newslett* 27:10-15.

It will be understood that various details of the invention can be changed without departing from the scope of the invention. Furthermore, the foregoing description is for the purpose of illustration only, and not for the purpose of limitation--the invention being defined by the claims.